Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 260 (2011) 30–75

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

A Strong Szegö–Widom Limit Theorem for operators with almost periodic diagonal

Torsten Ehrhardt [a,*,1], Steffen Roch [b], Bernd Silbermann [c]

[a] *Mathematics Department, University of California, Santa Cruz, CA 95064, USA*
[b] *Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstrasse 7, 64289 Darmstadt, Germany*
[c] *Technische Universität Chemnitz, Fakultät für Mathematik, 09107 Chemnitz, Germany*

Received 2 December 2009; accepted 26 July 2010

Communicated by D. Voiculescu

Abstract

The classical Strong Szegö–Widom Limit Theorem describes the asymptotic behavior of the determinants of the finite sections $P_n T(a) P_n$ of Toeplitz operators, i.e., of operators which have constant entries along each diagonal. We generalize these results to operators which have almost periodic sequences as their diagonals.

Keywords: Szegö–Widom Limit Theorems; Toeplitz operator; Almost Mathieu operator; Determinants of finite sections

1. Introduction

The classical results. The $n \times n$ Toeplitz matrices are defined as

$$T_n(a) = (a_{j-k}), \quad 0 \leqslant j, k \leqslant n-1,$$

* Corresponding author.
E-mail addresses: ehrhardt@math.ucsc.edu (T. Ehrhardt), roch@mathematik.tu-darmstadt.de (S. Roch), bernd.silbermann@mathematik.tu-chemnitz.de (B. Silbermann).

[1] Research supported in part by NSF Grant DMS-0901434.

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2010.07.016

where $a \in L^{\infty}(\mathbb{T})$ is a function defined on the unit circle $\mathbb{T} = \{z \in \mathbb{C}\colon |z| = 1\}$ with Fourier coefficients

$$a_k = \frac{1}{2\pi} \int_0^{2\pi} a\big(e^{ix}\big) e^{-ikx}\, dx, \quad k \in \mathbb{Z}.$$

Under certain assumptions on a, one of the several versions of the First Szegö Limit Theorem [23,26] states that

$$\lim_{n\to\infty} \frac{\det T_n(a)}{\det T_{n-1}(a)} = G[a],$$

while the Strong Szegö–Widom Limit Theorem [24], under additional assumptions, asserts that

$$\lim_{n\to\infty} \frac{\det T_n(a)}{G[a]^n} = E[a].$$

Therein $G[a]$ and $E[a]$ are well-defined and non-zero constants. At this point we will not discuss further details of the Szegö Limit Theorems, but refer to [8,9] and to Chapter 2 of [22], where also information about the long and rich history can be found. In [18] some milestones in this field are also mentioned.

Finite Toeplitz matrices $T_n(a)$ are the finite sections $P_n T(a) P_n$ of Toeplitz operators

$$T(a) = (a_{j-k}), \quad j, k \in \mathbb{Z}^+, \tag{1}$$

which act on $\ell^2(\mathbb{Z}^+)$, $\mathbb{Z}^+ := \{0, 1, 2, \ldots\}$. The projections P_n are defined by

$$P_n : \{x_0, x_1, x_2, \ldots\} \in \ell^2\big(\mathbb{Z}^+\big) \mapsto \{x_0, \ldots, x_{n-1}, 0, 0, \ldots\} \in \ell^2\big(\mathbb{Z}^+\big).$$

Furthermore, Toeplitz operators $T(a)$ arise as the compressions $PL(a)P$ of the Laurent operators

$$L(a) = (a_{j-k}), \quad j, k \in \mathbb{Z}, \tag{2}$$

which act on $\ell^2(\mathbb{Z})$. Here P is the Riesz projection operator

$$P : \{x_n\}_{n=-\infty}^{\infty} \in \ell^2(\mathbb{Z}) \mapsto \{y_n\}_{n=-\infty}^{\infty} \in \ell^2(\mathbb{Z}), \qquad y_n = \begin{cases} x_n & \text{if } n \geqslant 0, \\ 0 & \text{if } n < 0. \end{cases}$$

Laurent operators are constant on each diagonal. In other words, they are shift-invariant, i.e., they satisfy $U_{-n} L(a) U_n = L(a)$ for all $n \in \mathbb{Z}$, where U_n denotes the shift operator

$$U_n : \{x_k\}_{k=-\infty}^{\infty} \mapsto \{x_{k-n}\}_{k=-\infty}^{\infty} \tag{3}$$

acting on $\ell^2(\mathbb{Z})$.

The goal of this paper is to generalize the Strong Szegö–Widom Limit Theorem to finite sections $P_n P A P P_n$ in which A is a bounded linear operator acting on $\ell^2(\mathbb{Z})$ whose diagonals are almost periodic sequences. More precisely, our generalization will relate to so-called *band-dominated operators with almost periodic diagonals*. The paper [18], where two of the authors dealt with a generalization of the First Szegö Limit Theorem, is thus somewhat related.

Let us proceed with explaining the notion of (band-dominated) operators with almost periodic diagonals.

Operators with almost periodic diagonals. The set $AP(\mathbb{Z})$ of almost periodic sequences consists of all $a \in \ell^\infty(\mathbb{Z})$ for which the set

$$\{U_n a\colon\ n \in \mathbb{Z}\}$$

is relatively compact in the norm topology of $\ell^\infty(\mathbb{Z})$. Here

$$U_n : a \in \ell^\infty(\mathbb{Z}) \mapsto b \in \ell^\infty(\mathbb{Z}), \qquad b(k) := a(k-n), \tag{4}$$

is the shift operator acting (isometrically) on $\ell^\infty(\mathbb{Z})$. Clearly, the rules (3) and (4) are the same, only the underlying spaces are different. Moreover, we will prefer the notation $a(n)$ to the usual a_n for the entries of a sequence $a \in \ell^\infty(\mathbb{Z})$ because we will use the second notation for the Fourier coefficients of a sequence $a \in AP(\mathbb{Z})$ (see Section 2).

There is an equivalent definition of $AP(\mathbb{Z})$ as the closure in $\ell^\infty(\mathbb{Z})$ of the set of all finite linear combinations of sequences $\boldsymbol{e}_\xi \in \ell^\infty(\mathbb{Z})$, where

$$\boldsymbol{e}_\xi(n) = e^{2\pi i \xi n}, \quad n \in \mathbb{Z},\ \xi \in \mathbb{R}. \tag{5}$$

While a Laurent operator can be formally written as

$$L(a) = \sum_{n\in\mathbb{Z}} a^{(n)} U_n,$$

where $a^{(n)} \in \mathbb{C}$ are constants, operators with almost periodic diagonals can be formally written as

$$A = \sum_{n\in\mathbb{Z}} a^{(n)} U_n := \sum_{n\in\mathbb{Z}} \big(a^{(n)} I\big) U_n, \tag{6}$$

where $a^{(n)} \in AP(\mathbb{Z})$ and aI stands for the multiplication operator generated by $a \in \ell^\infty(\mathbb{Z})$,

$$aI : \{x_n\}_{n=-\infty}^{\infty} \in \ell^2(\mathbb{Z}) \mapsto \big\{a(n)x_n\big\}_{n=-\infty}^{\infty} \in \ell^2(\mathbb{Z}). \tag{7}$$

For sake of brevity, we will usually write $(a^{(n)} I)U_n$ as $a^{(n)} U_n$.

To make the notion of *operators with almost periodic diagonals* more precise, let $\mathcal{OAP}$ stand for the set of all bounded linear operators A on $\ell^2(\mathbb{Z})$ such that their n-th diagonal belongs to $AP(\mathbb{Z})$ for each $n \in \mathbb{Z}$. In other words,

$$D_n(A) := D(AU_{-n}) \in AP(\mathbb{Z}), \tag{8}$$

where $D(A) \in \ell^\infty(\mathbb{Z})$ stands for the main diagonal of a bounded linear operator A acting on $\ell^2(\mathbb{Z})$.

A subclass of $\mathcal{OAP}$ are *band-dominated operators* with almost periodic diagonals. This notion is used rather loosely. It is customary to have it referred to the closure of the set of *band operators* with almost periodic diagonals with respect to a suitable norm. Band operators are operators of the form (6) with the sum being finite. For some information on band-dominated operators see [17,18].

We are going to establish a generalization of the Strong Szegö–Widom Limit Theorem for certain subclasses of band-dominated operators with almost periodic diagonals. These classes are weighted Wiener type algebras and will be described later on in this introduction. However, in order to get some idea of what to expect as a generalization, let us look first at the case of block Toeplitz determinants.

Operators with periodic diagonals and block Toeplitz matrices. The class $AP(\mathbb{Z})$ contains all periodic sequences. An operator A acting on $\ell^2(\mathbb{Z})$ whose diagonals are periodic sequences with fixed period N can be identified with a block Laurent operator in an obvious way. Block Toeplitz and Laurent operators are defined by formulas (1) and (2), but with a being an $N \times N$ matrix-valued function, whose Fourier coefficients a_n are $N \times N$ matrices.

Finite block Toeplitz matrices are defined similarly. They arise from the finite sections of the afore-mentioned operators A by

$$T_n(a) = P_{nN} P A P P_{nN},$$

i.e., when the size of the finite section is a multiple of N.

The block case of the Strong Szegö Limit Theorem was established by Widom [26,9]. We are going to state this result for generating functions belonging to the Banach algebra $B = W \cap \mathrm{F}\ell^{2,2}_{1/2,1/2}$, which by definition consists of all (continuous) functions $a \in L^\infty(\mathbb{T})$ for which

$$\|a\|_B := \sum_{n=-\infty}^{\infty} |a_n| + \left(\sum_{n=-\infty}^{\infty} |n| \cdot |a_n|^2 \right)^{1/2} < \infty.$$

Theorem 1.1 *(Strong Szegö–Widom Limit Theorem). Let $a \in B^{N\times N}$, and assume that* $\det a(t) \neq 0$ *for all $t \in \mathbb{T}$ and that* $\det a(t)$ *has winding number zero. Then*

$$\lim_{n\to\infty} \frac{\det T_n(a)}{G[a]^n} = E[a] \tag{9}$$

where

$$G[a] = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log \det a\left(e^{ix}\right) dx \right)$$

and $E[a] = \det T(a) T(a^{-1})$.

The winding number condition is equivalent to the existence of a (continuous) logarithm of $\det a$ in B. Equivalently, we can require that a belongs to the connected component of the group of all invertible elements of $B^{N\times N}$ containing the unit element. Notice that this last condition is equivalent to requiring that the matrix function a is a finite product of exponentials of functions in $B^{N\times N}$ (see, e.g., [1,20]).

The constant $E[a]$ is defined in terms of an operator determinant. In fact, $T(a)T(a^{-1})$ equals identity plus a trace class operator. Only in the case $N=1$ (and perhaps in some other very special cases) a more explicit expression is known:

$$E[a]=\exp\left(\sum_{k=1}^{\infty}k(\log a)_k(\log a)_{-k}\right).$$

For more details on the Szegö–Widom Limit Theorem we refer to [8,9,12].

What to expect in the almost periodic case. In view of the results in the block Toeplitz case (i.e., in the case of periodic diagonals), we can now get an idea of what can and, maybe more important, what cannot be expected in the almost periodic case.

For instance, for general operators A with almost periodic diagonals one *cannot* expect an asymptotic formula of the kind

$$\lim_{n\to\infty}\frac{\det P_nPAPP_n}{G^n}=E. \tag{10}$$

Simple counterexamples can be constructed involving block diagonal Laurent operators.

What one can expect is that for certain strictly monotonically increasing sequences $h:\mathbb{Z}^+\to\mathbb{Z}^+$ we have

$$\lim_{n\to\infty}\frac{\det P_{h(n)}PAPP_{h(n)}}{G^{h(n)}}=E. \tag{11}$$

For instance, in the periodic case, we should take $h(n)=nN$. Then we have $G^N=G[a]$ in comparison with Theorem 1.1. In general such sequences h corresponds to the notion of a *distinguished sequence*, which plays a crucial role in this paper as it had in [18].

Moreover, we cannot expect an explicit expression for E besides that of an operator determinant.

Furthermore, we have to restrict the class of band-dominated operators. This corresponds to take the generating functions in Theorem 1.1 from the Banach algebra B instead of from $L^\infty(\mathbb{T})$.

Finally, in the almost periodic case a somewhat unexpected difficulty enters the scene. If one wants to obtain the asymptotics of determinants one should be able to analyze the asymptotics of traces in the first place. Indeed, to see that this is a necessary step consider the special case of $A=e^aI$ with $a\in AP(\mathbb{Z})$, and notice that $\det(P_nPAPP_n)=\exp(\operatorname{trace}(P_nPaIPP_n))$.

To compute the trace for the finite sections of $A\in\mathcal{OAP}$, let $a=D(A)\in AP(\mathbb{Z})$. Then

$$\operatorname{trace}(P_nPAPP_n)=\sum_{k=0}^{n-1}a(k). \tag{12}$$

It turns out that the right-hand side is $n \cdot M(a) + o(n)$, where $M(a)$ stands for the *mean* of an almost periodic sequence. Desirably we would like to have an error term $o(1)$. However, in the (almost) periodic case this cannot hold true. Yet one could expect that perhaps for distinguished sequences $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ we have

$$\sum_{k=0}^{h(n)-1} a(k) = h(n) \cdot M(a) + o(1), \quad n \to \infty. \tag{13}$$

For instance, in the case of a periodic sequence with period N, we can take $h(n) = Nn$. As will be shown by counterexamples in the last section of this paper, also this is not true in general! In the counterexamples the error term may not even be bounded.

Fortunately, for some classes of almost periodic sequences, the asymptotics (13) still holds true. The validity or failure is connected with the Fourier spectrum of the underlying sequence.

Illustration of the main results. The proof of the main results is very technical. Unfortunately, even to state the main results in full detail is quite technical, too. One reason is that one has to deal with three different objects:

1. additive subgroups Ξ of $\mathbb{R}/\mathbb{Z}$ which are related to the Fourier spectrum of almost periodic sequences;
2. Banach algebras $\mathcal{A}$ of almost periodic sequences involving a weight;
3. Banach algebras $\mathcal{R}$ of operators on $\ell^2(\mathbb{Z})$ whose diagonals are sequences of the previous Banach algebras.

The main result of this paper (Theorem 5.3) will be stated here in the introduction in full generality. It involves the just mentioned additive subgroups Ξ which we will only partially elaborate on here. Section 2 will provide full details on this matter. Sections 3–4 will provide the major steps towards the proof. For the proof it is necessary to introduce further auxiliary Banach algebras. Section 5 will give the main results with their proofs. Section 6 contains the counterexamples related to the asymptotics of the traces.

Let us now describe the main results. We start with an additive subgroup Ξ of $\mathbb{R}/\mathbb{Z}$, which at this point can be arbitrary. Here $\mathbb{R}/\mathbb{Z}$ denotes the additive group arising from $\mathbb{R}$ by identifying two numbers whose difference is an integer. As for notation, we will write ξ for both a real number and the corresponding equivalence class in $\mathbb{R}/\mathbb{Z}$.

An *admissible weight* $\boldsymbol{\beta}$ on Ξ is a function $\boldsymbol{\beta} : \Xi \to \mathbb{R}^+$ for which

$$1 \leqslant \boldsymbol{\beta}(\xi_1 + \xi_2) \leqslant \boldsymbol{\beta}(\xi_1)\boldsymbol{\beta}(\xi_2).$$

For instance, we can put $\boldsymbol{\beta}(\xi) = 1$. Another example of admissible weights can be constructed for finitely generated groups

$$\Xi = \{\xi = \alpha_1\xi_1 + \cdots + \alpha_N\xi_N\colon \alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{Z}^N\}$$

with given generators $(\xi_1, \ldots, \xi_N) \in \mathbb{R}^N$ by defining

$$\boldsymbol{\beta}(\xi) = \big(1 + |\alpha|\big)^\omega, \quad \omega \geqslant 0 \tag{14}$$

with $|\alpha| = \max|\alpha_i|$. Here we assume (for simplicity) that $1, \xi_1, \dots, \xi_N$ are linearly independent over the field $\mathbb{Q}$ to make sure that the coefficients α_k are uniquely determined by ξ.

For any such additive subgroup Ξ and any admissible weight $\boldsymbol{\beta}$ we define $\mathcal{A} = APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ as the set of all sequences $a \in \ell^\infty(\mathbb{Z})$ of the form

$$a = \sum_{\xi \in \Xi} a_\xi \boldsymbol{e}_\xi \quad \text{such that} \quad \|a\|_{\Xi,\boldsymbol{\beta}} := \sum_{\xi \in \Xi} \boldsymbol{\beta}(\xi)|a_\xi| < \infty$$

with $a_\xi \in \mathbb{C}$. We will show that $\mathcal{A}$ is a Banach algebra with the above norm which is continuously embedded in $AP(\mathbb{Z})$. All elements in $\mathcal{A}$ have Fourier spectrum contained in Ξ.

Next, given such a Banach algebra $\mathcal{A}$ and $\alpha_1, \alpha_2 \geqslant 0$ define $\mathcal{R} = \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ as the set of all operators A acting on $\ell^2(\mathbb{Z})$ which are of the form

$$A = \sum_{k \in \mathbb{Z}} a^{(k)} U_k, \quad a^{(k)} \in \mathcal{A}$$

such that

$$\|A\|_{\mathcal{R}} := \|a^{(0)}\|_{\mathcal{A}} + \sum_{k=1}^{\infty} \big((1+k)^{\alpha_1} \|a^{(k)}\|_{\mathcal{A}} + (1+k)^{\alpha_2} \|a^{(-k)}\|_{\mathcal{A}}\big) < \infty.$$

The last condition ensures that the defining series for A converges. The set $\mathcal{R}$ is a unital Banach algebra which is continuously embedded into $L(\ell^2(\mathbb{Z}))$ and contained in $\mathcal{OAP}$.

Finally, a strictly increasing sequence $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ is called *distinguished* for Ξ, if

$$\lim_{n \to \infty} e^{2\pi i \xi h(n)} = 1$$

for each $\xi \in \Xi$. Clearly, the previous condition is satisfied for all $\xi \in \Xi$ if it holds for a set of generators of the group Ξ. Moreover, one can show that all finitely generated groups Ξ possess distinguished sequences.

Now we can state an auxiliary result. It corresponds to a special case of Theorem 5.2.

Theorem 1.2. *Let Ξ, $\boldsymbol{\beta}$, $\mathcal{A}$, and $\mathcal{R}$ be as above, and assume that $\alpha_1 + \alpha_2 = 1$, $\alpha_1, \alpha_2 > 0$. Let $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ be a distinguished sequence for Ξ. Then, for $A_1, \dots, A_r \in \mathcal{R}$ and $A = e^{A_1} \cdots e^{A_r}$, we have*

$$\lim_{n \to \infty} \frac{\det(P_{h(n)} PAP P_{h(n)})}{\exp(\operatorname{trace}(P_{h(n)} P(A_1 + \cdots + A_r) P P_{h(n)}))} = \det\big(PAPA^{-1}P\big).$$

Therein the determinant on the right-hand side is a well-defined operator determinant of an operator acting on $\ell^2(\mathbb{Z}^+)$.

If we compare this theorem with the classical Strong Szegö–Widom Theorem (see Theorem 1.1), then the assumption that the operator A is a product of exponentials corresponds to the condition that the symbol of the block Toeplitz operator $T_n(a)$ has a non-vanishing determinant with winding number zero. In fact, we one can recover the classical theorem (although in the setting of a slightly different Banach algebra B) by considering $A = L(a)$, $h(n) = nN$, $\Xi = \{k/N\colon k \in \mathbb{Z}\}$, $\boldsymbol{\beta} \equiv 1$, and $\alpha_1 = \alpha_2 = 1/2$, where N is the block size.

The above theorem reduces the asymptotics of a determinant to the asymptotics of a trace. The computation of the trace seems easy. In fact, we have

$$\operatorname{trace}\bigl(P_{h(n)} P(A_1 + \cdots + A_r) P P_{h(n)}\bigr) = \sum_{k=0}^{h(n)-1} a(k)$$

where $a = D(A_1 + \cdots + A_r) \in \mathcal{A} = APW(\mathbb{Z}, \Xi, \boldsymbol{\beta}) \subseteq AP(\mathbb{Z})$ is the sequence arising from the main diagonals of the A_k's. However, as mentioned before, the form of the asymptotics that we would expect from the classical (periodic) case does not always hold in the almost periodic case. In order to ensure such an asymptotics, one needs to require a non-trivial extra assumption. This is an assumption on the underlying group Ξ and the weight $\boldsymbol{\beta}$.

We call the weight $\boldsymbol{\beta}$ *compatible* on Ξ if

$$\inf_{\xi \in \Xi,\, \xi \neq 0} \boldsymbol{\beta}(\xi) \cdot \|\xi\|_{\mathbb{R}/\mathbb{Z}} > 0$$

where $\|\xi\|_{\mathbb{R}/\mathbb{Z}} = \inf\{|\xi - n|\colon n \in \mathbb{Z}\}$. Now we have the following result.

Theorem 1.3. *Let $\boldsymbol{\beta}$ be an admissible and compatible weight on an additive subgroup Ξ of $\mathbb{R}/\mathbb{Z}$, and suppose that $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ is a distinguished sequence. Then, for $a \in APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$,*

$$\sum_{k=0}^{h(n)-1} a(k) = h(n) \cdot M(a) + o(1), \quad \textit{as } n \to \infty.$$

This theorem is stated and proved in Section 2 (see Theorem 2.7).

It is now clear that if we can combine the last two results, then we obtain

$$\lim_{n \to \infty} \frac{\det(P_{h(n)} P A P P_{h(n)})}{G^{h(n)}} = \det\bigl(P A P A^{-1} P\bigr) \quad \text{with } G = \exp\bigl(M(a)\bigr),$$

which is our main result (Theorem 5.3).

The main question which arises in connection with Theorem 1.3 is: *Given an additive subgroup Ξ, does there exist an admissible and compatible weight*? In addition, we do not want the weights to be growing too fast as this would restrict the classes of almost periodic sequences too much.

The counterexamples, which will be presented in Section 6, will imply that such compatible weight do not always exist, not even for every singly-generated group Ξ. On the other hand, there are positive results. They will be presented in Theorem 2.8 and in Section 2.3. First of all, it is shown that *almost every* finitely generated group Ξ possesses an admissible weight is of the form (14). Here "almost every" is understood in the sense of the Lebesgue's measure with respect to the generators of the group. To identify concrete groups which possess such weights is more complicated and in fact relies on deep results on diophantine approximation. For instance, positive results are obtained for groups generated by finitely many algebraic numbers.

Let us come back to the counterexamples. They concern a singly-generated group $\Xi = \{k\xi\colon k \in \mathbb{Z}\}$ with ξ being irrational. In fact, ξ must be a Liouville number. We will construct such numbers ξ (in terms of an infinite series) and corresponding almost periodic sequences

$$a = \sum_{k=1}^{\infty} a_k \boldsymbol{e}_{k\xi}$$

with exponentially decaying a_k's such that various properties are fulfilled. In particular, the following examples can be realized:

1. $a_k = b^{-k}$, $b > 1$, $x > 0$ arbitrary, and ξ suitably constructed: then there exists a distinguished sequence h such that

$$\sum_{k=0}^{h(n)-1} a(k) = x + o(1) \quad \text{as } n \to \infty.$$

2. $a_k = b^{-k}$, $b > 1$, $0 < \alpha < 1$ arbitrary, and ξ suitably constructed: then there exists a distinguished sequence h such that

$$\sum_{k=0}^{h(n)-1} a(k) = h(n)^{\alpha}\bigl(1 + o(1)\bigr) \quad \text{as } n \to \infty.$$

Notice that therein $M(a) = 0$, and observe that the asymptotics are different from those asserted in Theorem 1.3.

In view of the main result, yet other questions come up. How can one decide whether an operator is a product of exponentials (as is required in the assumptions of the main theorem)? We leave this important question open, although some comments are given at the end of Section 5. Furthermore, is there a way to compute the constant G directly in terms of A? Also this question is left open.

Outline of the paper and main idea of the proof. Let us give now a brief outline of the paper. Section 2 is devoted to almost periodic sequences. Besides elaborating on the notion of a distinguished sequence, we are going to define classes of almost sequences for which the asymptotics (13) holds. As mentioned before they are described by a certain condition, and we show, using known results from diophantine approximation, that this condition is in a certain sense generically fulfilled.

The proof our main result is based on a Banach algebra approach. This approach was introduced by one of the authors in [12] in the setting of the classical Strong Szegö–Widom Limit Theorem. The potential of this approach seems large and is by no means restricted to the classical results.

One could ask if the results of this paper could be proved by other methods. A recent approach to classical Strong Szegö–Widom Theorem is by means of the of Geronimo–Case–Borodin–Okounkov identity. However, it seems that for our operators this approach does not work.

The proof of the main result by the Banach algebra approach is done in Sections 3–5. In Section 3 the appropriate classes of operators on $\ell^2(\mathbb{Z})$ with almost periodic diagonals will be defined. The notions of suitability, rigidity, and shift-invariance are defined, too. The notion of suitability underlies the algebraic approach. In Section 4 certain Banach algebras of operators and of sequences are defined, which are derived from rigid, suitable and shift-invariant Banach algebras. These Banach algebras and their properties are the main ingredients to the Banach algebra approach. The main result is established in Section 5.

Finally, in Section 6 we present counterexamples to the asymptotics (13). These counterexamples contrast the positive results of Section 2 and show certain limitations of further generalizing the main results.

Our motivation for establishing the Szegö–Widom Limit Theorem for band-dominated operators with almost periodic coefficients comes from its most prominent example: the almost Mathieu operator

$$A = U_1 + aI + U_{-1}$$

Here $a(n) = \beta \sin(\xi n + \delta)$, and β, ξ, and δ are certain (real) constants. Our results would apply to the determinants of $P_n P(A - \lambda I) P P_n$ provided that the appropriate assumptions of the main theorem are fulfilled. For more information about almost periodic operators we refer to the monograph [6] and the more recent papers [2,16], where the long standing ten Martini problem is solved.

2. Banach algebras of almost periodic sequences

Let us start with some basic facts about $AP(\mathbb{Z})$. A basic reference for almost periodic functions (and thus also for almost periodic sequences) is [11].

For each $a \in AP(\mathbb{Z})$ its *mean* $M(a)$ is well defined by the limit

$$M(a) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a(k). \tag{15}$$

As already mentioned in the introduction, let $\mathbb{R}/\mathbb{Z}$ denote the additive group arising from $\mathbb{R}$ by identifying two numbers whose difference is an integer. We agree to denote a real number which is a representative of the equivalence class $\xi \in \mathbb{R}/\mathbb{Z}$ also by ξ.

The Fourier coefficients of a sequence $a \in AP(\mathbb{Z})$ are defined by

$$a_\xi = M(a\boldsymbol{e}_{-\xi}), \quad \xi \in \mathbb{R}/\mathbb{Z},$$

where $\boldsymbol{e}_\xi \in AP(\mathbb{Z})$ is the sequence defined in (5). The set of all $\xi \in \mathbb{R}/\mathbb{Z}$ for which $a_\xi \neq 0$ is called the Fourier spectrum of a. For each $a \in AP(\mathbb{Z})$ this set is at most countable.

A subclass of almost periodic sequences is the Wiener class $APW(\mathbb{Z})$, which is the set of all sequences $a \in AP(\mathbb{Z})$ with absolutely convergent Fourier series. In other words, it is the set of all sequences which can be represented as

$$a = \sum_\xi a_\xi \boldsymbol{e}_\xi \tag{16}$$

where the sum is taken over at most countably many elements $\xi \in \mathbb{R}/\mathbb{Z}$ and where

$$\|a\|_{APW(\mathbb{Z})} := \sum_\xi |a_\xi| < \infty.$$

The above norm makes $APW(\mathbb{Z})$ a Banach algebra, which is continuously embedded in $AP(\mathbb{Z})$.

The goal of this section is to provide information about the asymptotics of

$$\sum_{k=0}^{n-1} a(k) \tag{17}$$

as $n \to \infty$ for $a \in AP(\mathbb{Z})$. By the definition of the mean, this sum equals $nM(a) + o(n)$ as $n \to \infty$. In order to get a better estimate of the error term, let us consider the case where a is a *finite* sum (16). Using the obvious facts

$$M(\boldsymbol{e}_\xi) = \begin{cases} 1 & \text{if } \xi = 0, \\ 0 & \text{if } \xi \neq 0 \end{cases} \tag{18}$$

and

$$\sum_{k=0}^{n-1} \boldsymbol{e}_\xi(k) = \frac{1 - e^{2\pi i \xi n}}{1 - e^{2\pi i \xi}}, \tag{19}$$

it follows that

$$\sum_{k=0}^{n-1} a(k) = n \cdot M(a) + \sum_{\xi \neq 0} a_\xi \frac{1 - e^{2\pi i \xi n}}{1 - e^{2\pi i \xi}}.$$

The second term behaves oscillatory. Therefore, we are compelled to let run n not through all positive integers, but to restrict ourselves to certain strictly increasing sequence $h : \mathbb{Z}^+ \to \mathbb{Z}^+$. Indeed, if h is such that

$$\lim_{n \to \infty} e^{2\pi i \xi h(n)} = 1, \tag{20}$$

for all ξ in the Fourier spectrum of a, then

$$\sum_{k=0}^{h(n)-1} a(k) = h(n) \cdot M(a) + o(1), \quad n \to \infty.$$

At this point one is tempted to carry over this result to the case where a is an infinite sum (16), e.g., for general $a \in APW(\mathbb{Z})$. However, as will be shown in Section 6, this kind of asymptotics does not hold in general.

2.1. *Distinguished sequences for subalgebras of $AP(\mathbb{Z})$*

Throughout the rest of the paper, let $\mathcal{H}$ stand for the set of all strictly increasing sequences $h : \mathbb{Z}^+ \to \mathbb{Z}^+$.

All Banach algebras which we will consider have a unit element. When we say that $\mathcal{A}$ is a Banach subalgebra of a Banach algebra $\mathcal{B}$ we do not imply any relation between the norms of these two Banach algebras. We only require that the algebraic relations are compatible and that the unit elements are the same.

A Banach subalgebra $\mathcal{A}$ of $\ell^\infty(\mathbb{Z})$ will be called *shift-invariant* if, for each $a \in \mathcal{A}$ and $n \in \mathbb{Z}$, we have $U_n a \in \mathcal{A}$ and

$$\|U_n a\|_{\mathcal{A}} = \|a\|_{\mathcal{A}}. \tag{21}$$

Let $\mathcal{A}$ be a shift-invariant Banach subalgebra of $\ell^\infty(\mathbb{Z})$. A sequence $h \in \mathcal{H}$ will be called *distinguished for* $\mathcal{A}$ if for each $a \in \mathcal{A}$ we have

$$\lim_{n\to\infty} \|U_{-h(n)} a - a\|_{\mathcal{A}} = 0. \tag{22}$$

For certain classes of Banach algebras $\mathcal{A}$, the following proposition provides a simple criterion for a sequence h to be distinguished for $\mathcal{A}$.

Proposition 2.1. *Let Ξ be an additive subgroup of $\mathbb{R}/\mathbb{Z}$, and let $\mathcal{A}$ be a shift-invariant Banach subalgebra of $AP(\mathbb{Z})$ such that the linear span of*

$$\{\boldsymbol{e}_\xi\colon \xi \in \Xi\}$$

is contained and dense in $\mathcal{A}$. Then a sequence $h \in \mathcal{H}$ is distinguished for $\mathcal{A}$ if and only if for each $\xi \in \Xi$ we have

$$\lim_{n\to\infty} e^{2\pi i h(n)\xi} = 1. \tag{23}$$

Proof. From $U_{-h(n)}\boldsymbol{e}_\xi = e^{2\pi i h(n)\xi}\boldsymbol{e}_\xi$ it follows that

$$\|U_{-h(n)}\boldsymbol{e}_\xi - \boldsymbol{e}_\xi\|_{\mathcal{A}} = \left\|e^{2\pi i h(n)\xi}\boldsymbol{e}_\xi - \boldsymbol{e}_\xi\right\|_{\mathcal{A}} = \left|e^{2\pi i h(n)\xi} - 1\right| \cdot \|\boldsymbol{e}_\xi\|_{\mathcal{A}}.$$

This proves the "only if" part as well as the "if" part for $a = \boldsymbol{e}_\xi$. Moreover, assuming (23) it follows that (22) holds for finite linear combinations of $\boldsymbol{e}_\xi$. Using the density assumption an approximation argument implies the validity of (22) for all $a \in \mathcal{A}$. $\square$

The criterion can be relaxed further. In fact, it is sufficient to require (23) only for a set of generators of the group Ξ.

For the classes of Banach algebras to which the previous proposition can be applied, the question of h being distinguished is reduced to the underlying group Ξ. In such a setting we will say that $h \in \mathcal{H}$ is a *distinguished sequence for* Ξ.

Not for all additive subgroups Ξ there exist distinguished sequences. A necessary condition is that the Lebesgue measure of Ξ is zero, a sufficient condition is that Ξ is countable. (We will not prove these facts here.)

Example 2.2 *(Trivial example).* Let $N \in \mathbb{N}$ and

$$\Xi = \left\{\frac{k}{N} + \mathbb{Z}\colon k \in \mathbb{Z}\right\}.$$

Then $h(n) = Nn$ is a distinguished sequence for Ξ.

Example 2.3 *(Trivial example).* Let $\varXi = \mathbb{Q}/\mathbb{Z}$. Then $h(n) = n!$ is a distinguished sequence for $\varXi$.

Example 2.4. Let $\xi_1, \ldots, \xi_N \in \mathbb{R}$ be such that $\{1, \xi_1, \ldots, \xi_N\}$ are linearly independent over $\mathbb{Q}$, and let

$$\varXi = \{\alpha_1\xi_1 + \cdots + \alpha_N\xi_N\colon (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^N\}.$$

Then distinguished sequences exist. This follows from the well-known fact that the set

$$\{(n\xi_1, n\xi_2, \ldots, n\xi_N) \in (\mathbb{R}/\mathbb{Z})^N : n \in \mathbb{N}\}$$

is dense in $(\mathbb{R}/\mathbb{Z})^N$.

The next example is a special case of the previous one, where a particular distinguished sequence is described. For details, see [18, Example 2.10].

Example 2.5. Let $\varXi = \{\alpha\xi\colon \alpha \in \mathbb{Z}\}$ where $\xi \in \mathbb{R}$ is irrational. Let p_n/q_n be the n-th continued fraction of ξ. Then $h(n) = q_n$ is a distinguished sequence for $\varXi$. This follows from the well-known fact that $|q_n\xi - p_n| < 1/q_n$.

2.2. Weighted Wiener algebras of almost periodic sequences: the compatibility condition

In what follows we are going to define a concrete class of shift-invariant Banach algebras, namely weighted Wiener algebras of almost periodic sequences with their Fourier spectrum contained in a prescribed subgroup of $\mathbb{R}/\mathbb{Z}$.

Let $\varXi$ be an additive subgroup of $\mathbb{R}/\mathbb{Z}$. We call a mapping $\boldsymbol{\beta} : \varXi \to \mathbb{R}^+$ an *admissible weight on* $\varXi$ if

$$1 \leqslant \boldsymbol{\beta}(\xi_1 + \xi_2) \leqslant \boldsymbol{\beta}(\xi_1)\boldsymbol{\beta}(\xi_2) \tag{24}$$

for each $\xi_1, \xi_2 \in \varXi$. For such $\varXi$ and $\boldsymbol{\beta}$, let $APW(\mathbb{Z}, \varXi, \boldsymbol{\beta})$ stand for the set of all functions

$$a = \sum_{\xi\in\varXi} a_\xi \boldsymbol{e}_\xi \tag{25}$$

for which

$$\|a\|_{\varXi,\boldsymbol{\beta}} := \sum_{\xi\in\varXi} \boldsymbol{\beta}(\xi)|a_\xi| < \infty.$$

With the above norm, $APW(\mathbb{Z}, \varXi, \boldsymbol{\beta})$ becomes Banach space which is continuously embedded in $APW(\mathbb{Z})$. In particular, (25) is an absolutely convergent series. The following theorem guarantees that $APW(\mathbb{Z}, \varXi, \boldsymbol{\beta})$ is indeed a Banach algebra and that we can apply the results of the previous section.

Theorem 2.6. *Let $\boldsymbol{\beta}$ be an admissible weight on an additive subgroup $\varXi$ of $\mathbb{R}/\mathbb{Z}$. Then $APW(\mathbb{Z}, \varXi, \boldsymbol{\beta})$ is a shift-invariant, continuously embedded Banach subalgebra of $AP(\mathbb{Z})$, and the linear span of $\{\boldsymbol{e}_\xi\colon \xi \in \varXi\}$ is a dense subset.*

Proof. Let us first check that $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ is closed under multiplication. Assume that

$$a = \sum_{\xi \in \Xi} a_\xi \boldsymbol{e}_\xi, \qquad b = \sum_{\xi \in \Xi} b_\xi \boldsymbol{e}_\xi$$

belong to $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$. Since $\boldsymbol{e}_{\xi_1} \boldsymbol{e}_{\xi_2} = \boldsymbol{e}_{\xi_1+\xi_2}$, one has

$$ab = \sum_{\xi_1 \in \Xi} \sum_{\xi_2 \in \Xi} a_{\xi_1} b_{\xi_2} \boldsymbol{e}_{\xi_1+\xi_2} = \sum_{\xi \in \Xi} \left(\sum_{\xi_1 \in \Xi} a_{\xi_1} b_{\xi-\xi_1} \right) \boldsymbol{e}_\xi.$$

From the following estimate we can conclude that $ab \in APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ and also that $\|ab\| \leqslant \|a\| \|b\|$:

$$\begin{aligned}
\|ab\|_{\Xi,\boldsymbol{\beta}} &= \sum_{\xi \in \Xi} \boldsymbol{\beta}(\xi) \left| \sum_{\xi_1 \in \Xi} a_{\xi_1} b_{\xi-\xi_1} \right| \\
&\leqslant \sum_{\xi_1 \in \Xi} \sum_{\xi_2 \in \Xi} \boldsymbol{\beta}(\xi_1 + \xi_2) |a_{\xi_1} b_{\xi_2}| \\
&= \left(\sum_{\xi_1 \in \Xi} \boldsymbol{\beta}(\xi_1) |a_{\xi_1}| \right) \left(\sum_{\xi_2 \in \Xi} \boldsymbol{\beta}(\xi_2) |b_{\xi_2}| \right) \\
&= \|a\|_{\Xi,\boldsymbol{\beta}} \|b\|_{\Xi,\boldsymbol{\beta}}.
\end{aligned}$$

Notice that we have used that the weight $\boldsymbol{\beta}$ is admissible.

Let us now verify that $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ is shift-invariant. For $n \in \mathbb{Z}$ we have $U_{-n} \boldsymbol{e}_\xi = e^{2\pi i \xi n} \boldsymbol{e}_\xi$. Hence

$$a = \sum_{\xi \in \Xi} a_\xi \boldsymbol{e}_\xi \quad \text{implies} \quad U_{-n} a = \sum_{\xi \in \Xi} a_\xi e^{2\pi i \xi n} \boldsymbol{e}_\xi.$$

Because of $|e^{2\pi i \xi n}| = 1$ it follows that $U_{-n} a$ belongs to $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ whenever so does a, and one also has $\|U_{-n} a\|_{\Xi,\boldsymbol{\beta}} = \|a\|_{\Xi,\boldsymbol{\beta}}$.

The remaining statements are obvious. □

We are now going to establish the main theorem of this section. Therein we need to require a somewhat restrictive condition on Ξ and $\boldsymbol{\beta}$. In order to describe it, let us introduce

$$\|\xi\|_{\mathbb{R}/\mathbb{Z}} := \inf\bigl\{|\xi - n|: n \in \mathbb{Z}\bigr\}, \tag{26}$$

which corresponds to the natural metric on $\mathbb{R}/\mathbb{Z}$. Now we say that $\boldsymbol{\beta}$ is *compatible* on Ξ if

$$\inf_{\xi \in \Xi, \xi \neq 0} \boldsymbol{\beta}(\xi) \cdot \|\xi\|_{\mathbb{R}/\mathbb{Z}} > 0. \tag{27}$$

We will look in a few moments at the important question whether admissible and compatible weights exist.

Theorem 2.7. *Let $\boldsymbol{\beta}$ be an admissible and compatible weight on an additive subgroup Ξ of $\mathbb{R}/\mathbb{Z}$. If $a \in APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ and $h \in \mathcal{H}$ is a distinguished sequence for Ξ, then*

$$\sum_{k=0}^{h(n)-1} a(k) = h(n) \cdot M(a) + o(1), \quad n \to \infty. \tag{28}$$

Proof. First observe that (using $M(a) = a_0$)

$$\sum_{k=0}^{h(n)-1} a(k) - h(n)M(a) = \sum_{\xi \in \Xi,\, \xi \neq 0} a_\xi \left(\sum_{k=0}^{h(n)-1} \boldsymbol{e}_\xi(k) \right). \tag{29}$$

For (fixed) $\xi \in \Xi$, $\xi \neq 0$, the expression in the braces can be rewritten as follows,

$$\sum_{k=0}^{h(n)-1} \boldsymbol{e}_\xi(k) = \sum_{k=0}^{h(n)-1} e^{2\pi i \xi k} = \frac{1 - e^{2\pi i \xi h(n)}}{1 - e^{2\pi i \xi}},$$

and thus, by Proposition 2.1, it converges to zero as $n \to \infty$. On the other hand, the same expression can be estimated by

$$\left| \sum_{k=0}^{h(n)-1} \boldsymbol{e}_\xi(k) \right| = \frac{1}{|\sin(\pi\xi)|} \leqslant C\boldsymbol{\beta}(\xi),$$

where the constant C comes from the infimum (27). Hence, using dominated convergence, the expression (29) represents an absolutely convergent series which converges to zero as $n \to \infty$. □

Let us now turn to the question of the existence of admissible and compatible weights. We are not able to answer this question completely. In fact, it is connected to some deep question of diophantine approximation.

We are first going to show a pure existence result, namely, that for finitely generated subgroups there exists "generically" an admissible and compatible weight. In addition, the weight grows only polynomially. The basis of these observations is the following theorem, which is proved, e.g., in Section 3.5.3 ($N = 1$) and Section 4.3.2 ($N \geqslant 2$) of [5].

Theorem 2.8. *Let $N \geqslant 1$ and denote by $\mathcal{S}_N \subseteq (\mathbb{R}/\mathbb{Z})^N$ the set of all $(\xi_1, \dots, \xi_N)$ for which there exist $\omega > 0$ and $C > 0$ such that*

$$\|\alpha_1\xi_1 + \cdots + \alpha_N\xi_N\|_{\mathbb{R}/\mathbb{Z}} \geqslant C\Big(\max_{1 \leqslant i \leqslant N} |\alpha_i|\Big)^{-\omega} \tag{30}$$

for all $(\alpha_1, \dots, \alpha_N) \in \mathbb{Z}^N \setminus \{0\}$. Then the complement of $\mathcal{S}_N$ in $(\mathbb{R}/\mathbb{Z})^N$ has Hausdorff dimension $N - 1$, hence Lebesgue measure zero in $(\mathbb{R}/\mathbb{Z})^N$.

It is now obvious, that for $\xi \in \mathcal{S}_N$ we can define an admissible and compatible weight on

$$\Xi = \big\{\xi = \alpha_1\xi_1 + \cdots + \alpha_N\xi_N \in (\mathbb{R}/\mathbb{Z})^N : \alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}^N\big\}$$

by

$$\boldsymbol{\beta}(\xi) = \big(1 + |\alpha|\big)^{\omega}, \qquad |\alpha| = \max_{1 \leqslant i \leqslant N} |\alpha_i|.$$

In the case $N = 1$, the set $\mathcal{S}_1$ is just the set of all numbers which are *not* Liouville numbers.

2.3. Examples of Banach algebras $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$

The purpose of this section is to provide concrete examples of classes of additive subgroups Ξ and admissible and compatible weights $\boldsymbol{\beta}$, thus providing examples of Banach algebras $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ to which Theorem 2.7 can be applied.

Besides two trivial examples, we consider two other examples. In the first one, Ξ is finitely generated by algebraic numbers, in the second one, Ξ is generated by finitely many logarithms of algebraic numbers. Both examples rely on rather deep results from diophantine approximation.

Example 2.9 *(Trivial example).* Let $N \in \mathbb{N}$ and

$$\Xi = \left\{ \frac{k}{N} + \mathbb{Z}\colon k \in \mathbb{Z} \right\}.$$

Clearly, Ξ contains precisely N elements and we can define an admissible and compatible weight by $\boldsymbol{\beta}(\xi) = 1$. The Banach algebra $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ consists of all periodic sequences with period N.

Example 2.10 *(Trivial example).* Let $\Xi = \mathbb{Q}/\mathbb{Z}$. We can define an admissible and compatible weight by

$$\boldsymbol{\beta}(\xi) = q \quad \text{where } \xi = \frac{p}{q},\ p \in \mathbb{Z},\ q \in \mathbb{N},\ \text{and } p, q \text{ are co-prime.}$$

The Banach algebra $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ is the closure (with respect to the appropriate norm) of the set of all periodic sequences (with no conditions on the length of the period).

In order to establish non-trivial examples we need to resort to the theory of diophantine approximation. We refer to [4,10,25] as some references. Before stating these results, recall the notion of an algebraic number.

A number $\xi \in \mathbb{C}$ is called *algebraic* if there exists a (not identically vanishing) polynomial with integer coefficients that annihilates ξ. Among all such polynomials there exists one with smallest degree (or, equivalently, which is irreducible). The degree of the algebraic number is by definition the degree of this polynomial.

Theorem 2.11 *(Roth–Schmidt). Let $\xi_1, \ldots, \xi_N$ be algebraic numbers such that $\{1, \xi_1, \ldots, \xi_N\}$ is linearly independent over $\mathbb{Q}$. Then for every $d > 0$ there exists a constant $C_d > 0$ such that*

$$|\alpha_0 + \alpha_1 \xi_1 + \cdots + \alpha_N \xi_N| \geqslant C_d \Big(\max_{1 \leqslant i \leqslant N} |\alpha_i| \Big)^{-N-d}$$

for every $(\alpha_0, \alpha_1, \ldots, \alpha_N) \in \mathbb{Z}^{N+1}$.

This theorem was established by W.E. Schmidt [21]. In the special case $N = 1$ it is the celebrated theorem of Roth [19]. Roth's result was the culmination point of a development starting with Liouville's Theorem, where the exponent d depends on the degree of the algebraic number. The conditions on d were subsequently relaxed by A. Thue, C.L. Siegel, F.J. Dyson, A.O. Gelfond, and finally by K.F. Roth, who obtained, in some sense, the best possible result.

Notice that if $N = 1$ and if the (single) algebraic number ξ_1 has degree two, then Liouville's Theorem gives already a better result, namely the above inequality with $d = 0$.

Example 2.12. Let $\xi_1, \dots, \xi_N \in \mathbb{R}$ be *algebraic numbers* such that $\{1, \xi_1, \dots, \xi_N\}$ is linearly independent over $\mathbb{Q}$. Then every $\mathbb{Z}$-linear combination

$$\alpha_1\xi_1 + \cdots + \alpha_N\xi_N$$

can be considered as an element of $\mathbb{R}/\mathbb{Z}$ and determines the coefficients $\alpha_1, \dots, \alpha_N$ uniquely. Consider

$$\Xi = \{\alpha_1\xi_1 + \cdots + \alpha_N\xi_N\colon (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}^N\},$$

and for $\varepsilon > 0$ define the weight

$$\boldsymbol{\beta}(\xi) = (1 + |\alpha|)^{N+\varepsilon}, \qquad |\alpha| := \max\{|\alpha_1|, \dots, |\alpha_N|\}.$$

Obviously, $\boldsymbol{\beta}$ is admissible. The compatibility follows from the Roth–Schmidt Theorem.

For the next example we need to state two theorems as preparations. Let $\overline{\mathbb{Q}}$ denote the set of all algebraic numbers (possibly complex), and let

$$\Lambda = \{\lambda \in \mathbb{C}\colon \exp(\lambda) \in \overline{\mathbb{Q}}\}. \tag{31}$$

In other words, Λ is the set of logarithms of algebraic numbers.

Theorem 2.13 *(Baker). If $\lambda_1, \dots, \lambda_N \in \Lambda$ are linearly independent over $\mathbb{Q}$, then $1, \lambda_1, \dots, \lambda_N$ are linearly independent over $\overline{\mathbb{Q}}$.*

This theorem was established by Baker in 1966 [3]. Predecessors to this results were established by Gelfond, Kuzmin, and Gelfond–Schneider (see also [4,25]).

The next theorem is more general than the previous one. However, the previous theorem usually serves as a prerequisite for the next one.

Theorem 2.14 *(Feldman). Let $\lambda_1, \dots, \lambda_N \in \Lambda$ be linearly independent over $\mathbb{Q}$, and let $d \in \mathbb{N}$. Then there exist (effectively computable) constants $C, \omega > 0$ (depending only on $\lambda_1, \dots, \lambda_N$ and d) such that*

$$|\alpha_0 + \alpha_1\lambda_1 + \cdots + \alpha_N\lambda_N| \geqslant Ch^{-\omega} \tag{32}$$

for all $\alpha_0, \dots, \alpha_N \in \overline{\mathbb{Q}}$ of degree $\leqslant d$, where h is the maximum of the heights of $\alpha_0, \dots, \alpha_N$.

The height of an algebraic number $\alpha \neq 0$ is the maximum of the moduli of the coefficients of the annihilating irreducible polynomial with relatively prime integer coefficients. (The height of $\alpha = 0$ is zero by definition.)

Only the following two cases are of interest to us: if $\alpha \in \mathbb{Z}$, then its height equals $|\alpha|$, if $\alpha \in i\mathbb{Z}$, then its height equals $|\alpha|^2$.

The result was established by N.I. Feldman [13] in 1968 (although stated in a slightly different way). The constants, though explicit, are quite complicated. Improvements of those constants have been obtained subsequently. Lower estimates of non-algebraic type have been established earlier by A.O. Gelfond and A. Baker.

In the following example we will take generators $\xi_1, \dots, \xi_N$ for Ξ which belong either to $\Lambda \cap \mathbb{R}$ or $i\Lambda \cap \mathbb{R}$. The first class includes numbers such as

$$\log 5 \quad \text{or} \quad \log(1+\sqrt{2}).$$

The second class includes numbers such as

$$\pi = i\log(-1) \quad \text{or} \quad \arctan(\sqrt{2}) = i\log\frac{\sqrt{1-i\sqrt{2}}}{\sqrt{1+i\sqrt{2}}}.$$

Example 2.15. Let

$$\xi_1, \dots, \xi_M \in \Lambda \cap \mathbb{R}, \qquad \xi_{M+1}, \dots, \xi_N \in (i\Lambda \cap \mathbb{R}),$$

and assume that both $\{\xi_1, \dots, \xi_M\}$ and $\{\xi_{M+1}, \dots, \xi_N\}$ are linearly independent over $\mathbb{Q}$. Then the coefficients $\alpha_1, \dots, \alpha_N \in \mathbb{Z}$ of any linear combination $\alpha_1\xi_1 + \dots + \alpha_N\xi_N$ considered as an element in $\mathbb{R}/\mathbb{Z}$ are uniquely determined. Indeed, we can make an obvious substitution and write

$$\begin{aligned} &\alpha_0 + \alpha_1\xi_1 + \dots + \alpha_N\xi_N \\ &\quad = \alpha_0 + \alpha_1\lambda_1 + \dots + \alpha_M\lambda_M + (i\alpha_{M+1})\lambda_{M+1} + \dots + (i\alpha_N)\lambda_N \end{aligned} \tag{33}$$

with $\lambda_1, \dots, \lambda_N \in \Lambda$. It follows from the assumption that the set $\{\lambda_1, \dots, \lambda_N\}$ is linearly independent over $\mathbb{Q}$. Hence Theorem 2.13 implies that if the above linear combination is zero, then all their coefficients must be zero, too.

Now consider the subgroup

$$\Xi = \{\alpha_1\xi_1 + \dots + \alpha_N\xi_N\colon (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}^N\}$$

of $\mathbb{R}/\mathbb{Z}$. From Theorem 2.14 it follows that there exist constants $C > 0$ and ω such that estimate (32) holds, whence by using (33)

$$\|\alpha_1\xi_1 + \dots + \alpha_N\xi_N\|_{\mathbb{R}/\mathbb{Z}} \geqslant C\Big(\max_{0 \leqslant i \leqslant N} |\alpha_i|\Big)^{-2\omega}.$$

Note that the exponent 2ω comes from the fact that we have coefficients $i\alpha_k$, which have height $|\alpha_k|^2$. This implies by a simple computation that the weight

$$\boldsymbol{\beta}(\xi) = \big(1 + |\alpha|\big)^{2\omega}, \qquad |\alpha| := \max\big\{|\alpha_1|, \ldots, |\alpha_N|\big\}$$

is compatible and (obviously) admissible.

3. Banach algebras of operators on $\ell^2(\mathbb{Z})$

The purpose of this section is to find classes of Banach algebras of operators on $\ell^2(\mathbb{Z})$ for which a generalization of the Strong Szegö–Widom Limit Theorem can be established.

Our proof is based on a "Banach algebra approach". This approach was established in the case of classical Szegö–Widom Limit Theorem in [12]. The main idea is to reduce the asymptotics of determinants to the asymptotics of traces of certain operators, and it is accomplished by considering Banach algebras of "symbols", which are characterized by certain properties.

The main property that such Banach algebras have to possess is that of *suitability*. It generalizes the corresponding notion introduced first in [12].

There are also two other notions. They are trivial in the classical setting. The notion of *shift-invariance* of such a Banach algebra makes sure that the notion of a *distinguished sequence* can be considered properly. A third notion, that of *rigidity*, is of rather technical nature. One could probably work without it, but the changes necessary would make the presentation more tedious.

Before we start with this, let us introduce some notation. As usual, $L(H)$ stands for the Banach algebra of all bounded linear operators on a Hilbert space H. For $1 \leqslant p < \infty$, let $\mathcal{C}_p(H)$ refer to the Schatten–von Neumann class of operators on H, i.e., to the set of all compact operators $K \in L(H)$ for which

$$\|K\|_{\mathcal{C}_p} := \left(\sum_{n\geqslant 0} s_n(K)^p\right)^{1/p} < \infty$$

where $\{s_n(K)\}$ refers to the sequence of the decreasingly ordered non-zero eigenvalues of $(K^*K)^{1/2}$, with multiplicities taken into account.

Let P and J stand for the following operators on $\ell^2(\mathbb{Z})$,

$$P : \{x_n\}_{n=-\infty}^{\infty} \mapsto \{y_n\}_{n=-\infty}^{\infty} \quad \text{with } y_n = \begin{cases} x_n & \text{if } n \geqslant 0, \\ 0 & \text{if } n < 0, \end{cases}$$
$$J : \{x_n\}_{n=-\infty}^{\infty} \mapsto \{y_n\}_{n=-\infty}^{\infty} \quad \text{with } y_n = x_{-1-n},$$

and let $Q = I - P$. For each operator $A \in L(\ell^2(\mathbb{Z}))$ let

$$T(A) := PAP, \qquad H(A) := PAJP, \qquad \widetilde{A} := JAJ.$$

We will identify the image of the projection P with $\ell^2(\mathbb{Z}^+)$, and hence the operators $T(A)$ and $H(A)$ are to be considered as operators acting on $\ell^2(\mathbb{Z}^+)$. The identity operator on $\ell^2(\mathbb{Z}^+)$ will be denoted by P.

Using this notation, the identities

$$\begin{aligned} PABP &= PAPBP + PAQBP = PAPBP + PAJPJBP \\ &= (PAP)(PBP) + (PAJP)(PJBP), \end{aligned}$$

$$
\begin{aligned}
PABJP &= PAPBJP + PAQBJP = PAPBJP + PAJPJBJP \\
&= (PAP)(PBJP) + (PAJP)(PJBJP),
\end{aligned}
$$

read as

$$T(AB) = T(A)T(B) + H(A)H(\widetilde{B}), \tag{34}$$

$$H(AB) = T(A)H(B) + H(A)T(\widetilde{B}), \tag{35}$$

which reminds one on the basic identities between Toeplitz and Hankel operators. We will therefore call $T(A)$ and $H(A)$ the *Toeplitz* and the *Hankel operator* generated by $A \in L(\ell^2(\mathbb{Z}))$, respectively.

3.1. Rigid classes of operators on $\ell^2(\mathbb{Z})$

A set $\mathcal{R}$ of bounded linear operators on $\ell^2(\mathbb{Z})$ is called *rigid* if for each $A \in \mathcal{R}$ the following implication is true:

$$\text{If } T(A) \text{ or } T(\widetilde{A}) \text{ is compact, then } A = 0. \tag{36}$$

This notion is modeled after the corresponding property for Toeplitz operators. Clearly, each subset of a rigid set is rigid.

Theorem 3.1. *The class $\mathcal{OAP}$ is rigid.*

Proof. Because $A \in \mathcal{OAP}$ if and only if $\widetilde{A} \in \mathcal{OAP}$, it suffices to prove that PAP being compact implies $A = 0$.

Fix k and let $a^{(k)} := D_k(A) \in AP(\mathbb{Z})$. Write, formally,

$$PAP = \sum_{k \in \mathbb{Z}} P\big(a^{(k)} I\big) U_k P$$

to see that

$$D_k(PAP) = b^{(k)}$$

where

$$b^{(k)}(m) := \begin{cases} a^{(k)}(m) & \text{if } m \geqslant \max\{0, k\}, \\ 0 & \text{otherwise.} \end{cases}$$

We are going to use the peculiar property that for each almost periodic $a \in AP(\mathbb{Z})$ there exists a strictly increasing sequence $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that

$$\|U_{-h(n)} a - a\|_{\ell^\infty(\mathbb{Z})} \to 0, \quad \text{as } n \to \infty.$$

In particular, we have elementwise convergence, i.e., for each $m \in \mathbb{Z}$,

$$a(m) = \lim_{n\to\infty} a\big(m + h(n)\big).$$

We apply this to $a = a^{(k)}$ and observe that

$$D_k(U_{-h(n)} PAPU_{h(n)}) = U_{-h(n)} b^{(k)}.$$

The left-hand side converges to zero elementwise because of the compactness assumption, while the right-hand side converges to $a^{(k)}$. Hence $a^{(k)} = 0$. Since k was chosen arbitrarily we can conclude that $A = 0$. □

3.2. *Suitable Banach algebras of operators on $\ell^2(\mathbb{Z})$*

A Banach algebra $\mathcal{R}$ of bounded linear operators on $\ell^2(\mathbb{Z})$ will be called *suitable* if the following conditions are satisfied:

(a) $\mathcal{R}$ is continuously embedded into $L(\ell^2(\mathbb{Z}))$ and

$$\|A\|_{L(\ell^2(\mathbb{Z}))} \leqslant \|A\|_{\mathcal{R}} \quad \text{for all } A \in \mathcal{R},$$

(b) For all $A, B \in \mathcal{R}$, the operators $PAQBP$ and $QAPBQ$ are trace class, and there is a constant M such that

$$\max\big\{\|PAQBP\|_{\mathcal{C}_1}, \|QAPBQ\|_{\mathcal{C}_1}\big\} \leqslant M\|A\|_{\mathcal{R}}\|B\|_{\mathcal{R}} \quad \text{for all } A, B \in \mathcal{R}.$$

We proceed with establishing some basic observations, which can easily be verified.

(i) Condition (b) can be reformulated by saying that both $H(A)H(\widetilde{B})$ and $H(\widetilde{A})H(B)$ are trace class and

$$\max\big\{\big\|H(A)H(\widetilde{B})\big\|_{\mathcal{C}_1}, \big\|H(\widetilde{A})H(B)\big\|_{\mathcal{C}_1}\big\} \leqslant M\|A\|_{\mathcal{R}}\|B\|_{\mathcal{R}}$$

for all $A, B \in \mathcal{R}$. Moreover, by induction, condition (b) implies that

$$P(\Pi A_i)P - \Pi(PA_iP) \in \mathcal{C}_1\big(\ell^2(\mathbb{Z})\big) \tag{37}$$

for each finite product of elements $A_i \in \mathcal{R}$.

(ii) The set of all operators of the form $PAP + QBQ$ with $A, B \in L(\ell^2(\mathbb{Z}))$ equipped with the norm of $L(\ell^2(\mathbb{Z}))$ forms a suitable Banach algebra.

(iii) If $\mathcal{R}$ is a suitable Banach algebra with respect to the norm $\|.\|_{\mathcal{R}}$, then the set $\widetilde{\mathcal{R}}$ of all operators $\widetilde{A} = JAJ$ with $A \in \mathcal{R}$ forms a suitable Banach algebra with respect to the norm $\|JAJ\|_{\widetilde{\mathcal{R}}} := \|A\|_{\mathcal{R}}$.

(iv) If $\mathcal{R}$ is a suitable Banach algebra and $\mathcal{A}$ is another Banach algebra which is continuously embedded in $L(\ell^2(\mathbb{Z}))$, then $\mathcal{R} \cap \mathcal{A}$ is a suitable Banach algebra with the norm $\|A\|_{\mathcal{R}\cap\mathcal{A}} := \|A\|_{\mathcal{R}} + \|A\|_{\mathcal{A}}$.

Another class of elementary examples (besides the one mentioned in (ii)) of suitable Banach algebras is the following.

Example 3.2 *(Operators with Schatten class corners).* For $p, q \geqslant 1$, let

$$\mathcal{R}_{p,q} := \{A \in L(\ell^2(\mathbb{Z})) : PAQ \in \mathcal{C}_p(\ell^2(\mathbb{Z})),\ QAP \in \mathcal{C}_q(\ell^2(\mathbb{Z}))\},$$

and introduce a norm by

$$\|A\|_{\mathcal{R}_{p,q}} := \|A\|_{L(\ell^2(\mathbb{Z}))} + \|PAQ\|_{\mathcal{C}_p} + \|QAP\|_{\mathcal{C}_q}. \tag{38}$$

Proposition 3.3. *$\mathcal{R}_{p,q}$ is a Banach algebra with the norm* (38) *if $p, q \geqslant 1$, and it is a suitable Banach algebra if, in addition, $1/p + 1/q = 1$.*

Proof. The proof follows easily from the fact that $\mathcal{C}_p$ is a closed ideal of $L(\ell^2(\mathbb{Z}))$ under the norm $\|.\|_{\mathcal{C}_p}$ and from $\mathcal{C}_p\mathcal{C}_q \subseteq \mathcal{C}_1$ whenever $1/p + 1/q = 1$, in which case one also has

$$\|AB\|_{\mathcal{C}_1} \leqslant \|A\|_{\mathcal{C}_p}\|B\|_{\mathcal{C}_q}$$

for arbitrary operators $A \in \mathcal{C}_p$ and $B \in \mathcal{C}_q$. □

Before giving our main example, let us introduce the notions of shift-invariance and of a distinguished sequence, this time for Banach algebras of operators on $\ell^2(\mathbb{Z})$.

A Banach algebras $\mathcal{R}$ of bounded linear operators on $\ell^2(\mathbb{Z})$ will be called *shift-invariant* if for each $A \in \mathcal{R}$ and $n \in \mathbb{Z}$ we have

$$U_{-n}AU_n \in \mathcal{R} \quad \text{and} \quad \|U_{-n}AU_n\|_{\mathcal{R}} = \|A\|_{\mathcal{R}}. \tag{39}$$

Let $\mathcal{R}$ be a shift-invariant Banach algebra of operators on $\ell^2(\mathbb{Z})$. A sequence $h \in \mathcal{H}$ will be called *distinguished for $\mathcal{R}$*, if for each $A \in \mathcal{R}$,

$$\lim_{n\to\infty} \|U_{-h(n)}AU_{h(n)} - A\|_{\mathcal{R}} = 0. \tag{40}$$

Recall that $\mathcal{H}$ denotes the set of all strictly increasing sequences $h : \mathbb{Z}^+ \to \mathbb{Z}^+$.

This definition of shift-invariance for Banach subalgebras of $L(\ell^2(\mathbb{Z}))$ generalizes the corresponding notion for Banach subalgebras of $\ell^\infty(\mathbb{Z})$ considered in the Section 2.1. Indeed, each element in $\ell^\infty(\mathbb{Z})$ can be understood as a diagonal operator acting on $\ell^2(\mathbb{Z})$.

Example 3.4 *(A generalized Wiener algebra).* Given $\alpha_1, \alpha_2 \geqslant 0$ define a weight function $\boldsymbol{\alpha}$ on $\mathbb{Z}$ by

$$\boldsymbol{\alpha}(k) = \begin{cases} (1+k)^{\alpha_1} & \text{if } k \geqslant 0, \\ (1+|k|)^{\alpha_2} & \text{if } k < 0. \end{cases} \tag{41}$$

Given, in addition, a shift-invariant Banach subalgebra $\mathcal{A}$ of $\ell^\infty(\mathbb{Z})$, let $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ stand for the set all operators $A \in L(\ell^2(\mathbb{Z}))$ for which

$$D_k(A) \in \mathcal{A}$$

for each $k \in \mathbb{Z}$ and

$$\|A\|_{\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})} := \sum_{k\in\mathbb{Z}} \boldsymbol{\alpha}(k) \big\| D_k(A) \big\|_{\mathcal{A}} < \infty. \tag{42}$$

Each operator in $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ can be written as

$$A = \sum_{k\in\mathbb{Z}} D_k(A)U_k,$$

where the series converges absolutely in the norm of $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ (hence in the norm of $L(\ell^2(\mathbb{Z}))$).

Conversely, given any collection of $a^{(k)} \in \mathcal{A}$, $k \in \mathbb{Z}$, satisfying

$$\sum_{k\in\mathbb{Z}} \boldsymbol{\alpha}(k) \big\| a^{(k)} \big\|_{\mathcal{A}} < \infty$$

a corresponding operator $A \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is defined by the absolutely convergent series

$$A = \sum_{k\in\mathbb{Z}} \big(a^{(k)} I\big) U_k, \tag{43}$$

where $a^{(k)} I$ is the multiplication operator defined in (7).

The main properties of $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ are stated in the following theorem.

Theorem 3.5. *Let $\alpha_1, \alpha_2 \geqslant 0$, and let $\mathcal{A}$ be a shift-invariant Banach subalgebra of $\ell^\infty(\mathbb{Z})$. Then $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is a shift-invariant Banach subalgebra of $L(\ell^2(\mathbb{Z}))$. Moreover,*

(a) *if $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$, then $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is suitable,*
(b) *if $\mathcal{A} \subseteq AP(\mathbb{Z})$, then $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is rigid.*

Proof. It is easy to verify that $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is a Banach space. Because of the estimate

$$\begin{aligned}
\|A\|_{L(\ell^2(\mathbb{Z}))} &= \Big\| \sum_{k\in\mathbb{Z}} D_k(A)U_k \Big\|_{L(\ell^2(\mathbb{Z}))} \\
&\leqslant \sum_{k\in\mathbb{Z}} \big\| D_k(A) \big\|_{\ell^\infty(\mathbb{Z})} \\
&\leqslant \|A\|_{\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})},
\end{aligned}$$

$\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is continuously embedded in $L(\ell^2(\mathbb{Z}))$. Now let A be given by (43). Then

$$U_{-n} A U_n = \sum_{k\in\mathbb{Z}} \big(U_{-n} a^{(k)}\big) U_k.$$

From this and the shift-invariance of $\mathcal{A}$ it is easy to conclude that $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is shift-invariant.

In order to verify that $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is closed under multiplication, let

$$A = \sum_{k\in\mathbb{Z}} a^{(k)} U_k, \qquad B = \sum_{k\in\mathbb{Z}} b^{(k)} U_k$$

be elements of $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$. Then

$$\begin{aligned} AB &= \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} a^{(k_1)} U_{k_1} b^{(k_2)} U_{k_2} \\ &= \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} a^{(k_1)} \big(U_{k_1} b^{(k_2)}\big) U_{k_1+k_2} \\ &= \sum_{n\in\mathbb{Z}} c^{(n)} U_n \end{aligned}$$

with

$$c^{(n)} = \sum_{k\in\mathbb{Z}} a^{(k)} \big(U_k b^{(n-k)}\big).$$

The fact that $c^{(n)} \in \mathcal{A}$ can be seen from the estimate

$$\begin{aligned} \sum_{n\in\mathbb{Z}} \boldsymbol{\alpha}(n) \big\|c^{(n)}\big\|_{\mathcal{A}} &\leqslant \sum_{n,k\in\mathbb{Z}} \boldsymbol{\alpha}(k)\boldsymbol{\alpha}(n-k) \big\|a^{(k)}\big\|_{\mathcal{A}} \big\|U_k b^{(n-k)}\big\|_{\mathcal{A}} \\ &= \|A\|_{\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})} \|B\|_{\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})}, \end{aligned}$$

which also implies that $AB \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ and $\|AB\| \leqslant \|A\| \cdot \|B\|$. Here we have used the assumption that $\mathcal{A}$ is shift-invariant and that

$$\boldsymbol{\alpha}(n) \leqslant \boldsymbol{\alpha}(k)\boldsymbol{\alpha}(n-k),$$

which in turn can be verified straightforwardly.

(a): We show that if $\alpha_2 > 0$ and $A \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$, then $PAQ \in \mathcal{C}_p$ with $p = 1/\alpha_2$. Indeed, write

$$A = \sum_{k\in\mathbb{Z}} D_k(A) U_k.$$

Because P commutes with the diagonal operators $D_k(A)$ we then have

$$\begin{aligned} PAQ &= \sum_{k\in\mathbb{Z}} D_k(A) P U_k Q \\ &= \sum_{k>0} D_k(A) P U_k Q. \end{aligned}$$

The singular values of PU_kQ are just the square roots of the eigenvalues of $PU_kQU_{-k}P = P_k$. Thus, 1 is a k-fold singular value of PU_kQ, and it is the only non-zero singular value of that operator. Hence, $\|PU_kQ\|_{\mathcal{C}_p} = k^{1/p} = k^{\alpha_2}$, which implies the norm estimate

$$\|PAQ\|_{\mathcal{C}_p} \leqslant \sum_{k>0} \|D_k(A)\|_{\ell^\infty(\mathbb{Z})} \|PU_kQ\|_{\mathcal{C}_p} \leqslant \sum_{k>0} k^{\alpha_1} \|D_k(A)\|_{\mathcal{A}} < \infty.$$

In particular, $PAQ \in \mathcal{C}_p$. Similarly we obtain $QAP \in \mathcal{C}_q$ with $1/q = \alpha_1$ whenever $A \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ and $\alpha_1 > 0$.

Consequently, if $\alpha_1 + \alpha_2 = 1$, then $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is continuously embedded into the Banach algebra $\mathcal{R}_{p,q}$ with Schatten class corners as defined in Example 3.2 where $p = 1/\alpha_2$ and $q = 1/\alpha_1$, and one has

$$\|A\|_{\mathcal{R}_{p,q}} \leqslant \|A\|_{\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})}$$

for all operators $A \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$. The remark made in (iv) above yields finally that $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ is a suitable Banach algebra.

(b): If $\mathcal{A} \subseteq AP(\mathbb{Z})$, then $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A}) \subseteq \mathcal{OAP}$. Now Theorem 3.1 applies. □

Proposition 3.6. *Let $\mathcal{A}$ be shift-invariant Banach subalgebra of $\ell^\infty(\mathbb{Z})$, and let $\alpha_1, \alpha_2 \geqslant 0$. Then $h \in \mathcal{H}$ is distinguished for $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ if and only if h is distinguished for $\mathcal{A}$.*

Proof. The "only if" part follows by considering the special case of $A = aI \in \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ with $a \in \mathcal{A}$.

As to the "if" part, consider A of the form (43) and assume $U_{-h(n)}a^{(k)} \to a^{(k)}$ in the norm of $\mathcal{A}$ for each $k \in \mathbb{Z}$. Because

$$U_{-h(n)}\left(\sum_{k\in\mathbb{Z}} a^{(k)}U_k\right)U_{h(n)} = \sum_{k\in\mathbb{Z}}\big(U_{-h(n)}a^{(k)}\big)U_k$$

we can conclude that $U_{-h(n)}AU_{h(n)} \to A$ in the norm of $\mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$. Indeed, we use the fact that the operators $U_{-h(n)}$ are isometries on $\mathcal{A}$ in order to apply dominated converges of the infinite series. □

The Banach algebra $\mathcal{R} = \mathcal{W}_{\alpha_1,\alpha_2}(\mathcal{A})$ with $\mathcal{A} = APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ is the algebra of operators for which we will state later our generalization of the Strong Szegö–Widom Limit Theorem. Of course, we also need to assume that $\boldsymbol{\beta}$ is admissible and compatible on Ξ.

As we have seen, $\mathcal{R}$ satisfies the following conditions which one should keep in mind: it is a Banach algebra, it is suitable, rigid, and shift-invariant. It also contains all band operators with diagonals in $APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$. There are certainly other algebras with these properties which could take the place of $\mathcal{R}$. Generalizations of the Krein and the Wiener–Krein algebras are good candidates (see Chapter 5 in [7] and [12,15]). We will not pursue this topic further since the main applications involve band operators with almost periodic coefficients which are already covered by this Banach algebra.

3.3. An operator determinant identity

The goal of this subsection is to generalize an operator determinant identity which was established in [12] in the context of suitable Banach algebras of Laurent operators (or, functions on the unit circle, if we stick to the original setting in [12]). We show that this identity holds for suitable Banach algebras in our context, i.e., it holds also in the non-commutative setting.

In what follows we employ the notion of an analytic Banach algebra valued function. We refer to [20] for a definition and basic properties.

Proposition 3.7. *Let $\mathcal{R}$ be a suitable Banach algebra, and let $A_1, \dots, A_r \in \mathcal{R}$. Then the functions*

$$F_0(\lambda_1, \dots, \lambda_r) := T\big(e^{\lambda_1 A_1} \cdots e^{\lambda_r A_r}\big) \cdot e^{-\lambda_r T(A_r)} \cdots e^{-\lambda_1 T(A_1)} - P,$$
$$F_1(\lambda_1, \dots, \lambda_r) := T\big(e^{\lambda_1 \widetilde{A}_1} \cdots e^{\lambda_r \widetilde{A}_r}\big) \cdot e^{-\lambda_r T(\widetilde{A}_r)} \cdots e^{-\lambda_1 T(\widetilde{A}_1)} - P$$

are analytic with respect to each variable $\lambda_k \in \mathbb{C}$ and take values in the ideal of the trace class operators on $\ell^2(\mathbb{Z}^+)$.

Proof. The analyticity follows easily since $\lambda \mapsto e^{\lambda B}$ is an analytic function for each bounded linear operator B and since the product of analytic functions is analytic again. To get the trace class property of F_0 note that

$$F_0(\lambda_1, \dots, \lambda_r) - T\big(e^{\lambda_1 A_1}\big) \cdots T\big(e^{\lambda_r A_r}\big) \cdot e^{-\lambda_r T(A_r)} \cdots e^{-\lambda_1 T(A_1)} + P$$

is trace class by (37). Thus, it is sufficient to prove that

$$T\big(e^{\lambda A}\big) e^{-\lambda T(A)} - P \in \mathcal{C}_1 \tag{44}$$

for every $A \in \mathcal{R}$. Again by (37), one has

$$T\big(A^k\big) = P A^k P = (PAP)^k + R_k \quad \text{for each } k \in \mathbb{Z}^+$$

with $R_k \in \mathcal{C}_1$. This implies that

$$T\big(e^{\lambda A}\big) = \sum_{k \geqslant 0} \frac{\lambda^k}{k!} T\big(A^k\big) = \sum_{k \geqslant 0} \frac{\lambda^k}{k!} \big((PAP)^k + R_k\big) = e^{\lambda T(A)} + \sum_{k \geqslant 0} \frac{\lambda^k}{k!} R_k.$$

We claim that the latter series converges in $\mathcal{C}_1$. Indeed, by properties (a) and (b) of a suitable Banach algebra,

$$\begin{aligned}
\|R_k\|_{\mathcal{C}_1} &= \big\| P A^k P - (PAP)^k \big\|_{\mathcal{C}_1} \\
&= \big\| PAPA^{k-1} P - (PAP)^k \big\|_{\mathcal{C}_1} + \big\| PAQA^{k-1} P \big\|_{\mathcal{C}_1} \\
&\leqslant \|PAP\|_{L(\ell^2(\mathbb{Z}^+))} \big\| PA^{k-1} P - (PAP)^{k-1} \big\|_{\mathcal{C}_1} + \|A\|_{\mathcal{R}} \big\| A^{k-1} \big\|_{\mathcal{R}} \\
&\leqslant \|A\|_{\mathcal{R}} \big\| PA^{k-1} P - (PAP)^{k-1} \big\|_{\mathcal{C}_1} + \|A\|_{\mathcal{R}}^k.
\end{aligned}$$

Repeating these arguments r times yields

$$\|R_k\|_{\mathcal{C}_1} \leqslant \|A\|_{\mathcal{R}}^r \left\| PA^{k-r}P - (PAP)^{k-r} \right\|_{\mathcal{C}_1} + r\|A\|_{\mathcal{R}}^k$$

for each $r = 1, \dots, k$. In particular, $\|R_k\|_{\mathcal{C}_1} \leqslant k\|A\|_{\mathcal{R}}^k$. Thus, the series

$$\sum_{k\geqslant 1} \frac{|\lambda|^k}{k!} k\|A\|_{\mathcal{R}}^k = \sum_{k\geqslant 1} \frac{|\lambda|^k}{(k-1)!} \|A\|_{\mathcal{R}}^k = \|\lambda A\|_{\mathcal{R}} \sum_{k\geqslant 0} \frac{\|\lambda A\|_{\mathcal{R}}^k}{k!}$$

is a convergent majorant for $\sum_{k\geqslant 0} \frac{\lambda^k}{k!} R_k$ whence the convergence of the latter series in $\mathcal{C}_1$. This settles the proof for F_0. The proof for F_1 follows from that for F_0 if one takes into account that $J\mathcal{R}J$ is a suitable Banach algebra whenever $\mathcal{R}$ is so (see the remark made in (iii) at the beginning of Section 3.2). $\square$

The preceding proposition implies that

$$\det T\left(e^{\lambda_1 A_1} \cdots e^{\lambda_r A_r}\right) \cdot e^{-\lambda_r T(A_r)} \cdots e^{-\lambda_1 T(A_1)},$$
$$\det T\left(e^{\lambda_1 \widetilde{A}_1} \cdots e^{\lambda_r \widetilde{A}_r}\right) \cdot e^{-\lambda_r T(\widetilde{A}_r)} \cdots e^{-\lambda_1 T(\widetilde{A}_1)}$$

are well-defined operator determinants which depend analytically on each of the complex parameters λ_k. The above determinants has to be understood as operator determinants. For basic information about operator determinants and operator traces we refer to [14].

The announced operator determinant identity reads as follows.

Theorem 3.8. *Let $\mathcal{R}$ be a suitable Banach algebra, and let $A_1, \dots, A_r \in \mathcal{R}$. Then for each $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ the operator determinant*

$$f(\lambda_1, \dots, \lambda_r) := \det T\left(e^{\lambda_1 \widetilde{A}_1} \cdots e^{\lambda_r \widetilde{A}_r}\right) \cdot e^{-\lambda_r T(\widetilde{A}_r)} \cdots e^{-\lambda_1 T(\widetilde{A}_1)},$$

is equal to the operator determinant

$$g(\lambda_1, \dots, \lambda_r) := \det e^{\lambda_1 T(A_1)} \cdots e^{\lambda_r T(A_r)} \cdot T\left(e^{-\lambda_r A_r} \cdots e^{-\lambda_1 A_1}\right).$$

For the proof of this theorem, we will further need the following auxiliary facts.

Fact 1. Let A, B be bounded linear operators on a separable Hilbert space H such that both AB and BA are trace class operators. Then AB and BA have the same trace, and $I + AB$ and $I + BA$ have the same determinant.

Fact 2. Let $U \subseteq \mathbb{C}$ be a domain, and let $K : U \to \mathcal{C}_1$ be an analytic function. Assume further that $A(\lambda) := I + K(\lambda)$ is invertible (or, equivalently, that $\det A(\lambda) \neq 0$) for every $\lambda \in U$. Then

$$\frac{(\det A(\lambda))'}{\det A(\lambda)} = \operatorname{trace}\left(A'(\lambda)A^{-1}(\lambda)\right) = \operatorname{trace}\left(A^{-1}(\lambda)A'(\lambda)\right)$$

for every $\lambda \in U$.

Proof of Theorem 3.8. Set $A := e^{\lambda_1 A_1} \cdots e^{\lambda_r A_r}$. Due to analyticity, it suffices to prove the equality of f and g for parameters $\lambda_1, \ldots, \lambda_r$ which are close to zero. For these parameters, the operators

$$\widetilde{A} = JAJ = e^{\lambda_1 JA_1J} \cdots e^{\lambda_r JA_rJ} = e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}$$

and

$$A^{-1} = e^{-\lambda_r A_r} \cdots e^{-\lambda_1 A_1}$$

are close to the identity, and the corresponding Toeplitz operators $T(\widetilde{A}) = P\widetilde{A}P$ and $T(A^{-1}) = PA^{-1}P$ are invertible on $\ell^2(\mathbb{Z}^+)$. Hence, f and g are non-zero.

The equality $f(\lambda_1, \ldots, \lambda_r) = g(\lambda_1, \ldots, \lambda_r)$ will be proved by induction on r. For $r = 0$, the assertion is trivial. Assume the equality is already proved for $r-1$, e.g., assume that $f(\lambda_1, \ldots, \lambda_{r-1}, 0) = g(\lambda_1, \ldots, \lambda_{r-1}, 0)$ for all parameters sufficiently close to zero. For fixed $\lambda_1, \ldots, \lambda_{r-1}$, let

$$B(\lambda_r) := T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}\bigr) \cdot e^{-\lambda_r T(\widetilde{A_r})} \cdots e^{-\lambda_1 T(\widetilde{A_1})}.$$

By the product rule of differentiation,

$$\begin{aligned} B'(\lambda_r) = {} & T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}} \widetilde{A_r}\bigr) \cdot e^{-\lambda_r T(\widetilde{A_r})} \cdots e^{-\lambda_1 T(\widetilde{A_1})} \\ & - T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}\bigr) \cdot T(\widetilde{A_r}) e^{-\lambda_r T(\widetilde{A_r})} \cdots e^{-\lambda_1 T(\widetilde{A_1})} \end{aligned}$$

whence

$$\begin{aligned} B'(\lambda_r)B^{-1}(\lambda_r) &= T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}} \widetilde{A_r}\bigr) T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}\bigr)^{-1} \\ &\quad - T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}\bigr) T(\widetilde{A_r}) T\bigl(e^{\lambda_1 \widetilde{A_1}} \cdots e^{\lambda_r \widetilde{A_r}}\bigr)^{-1} \\ &= T(\widetilde{A}\widetilde{A_r}) T(\widetilde{A})^{-1} - T(\widetilde{A}) T(\widetilde{A_r}) T(\widetilde{A})^{-1} \\ &= H(\widetilde{A}) H(A_r) T(\widetilde{A})^{-1} \end{aligned}$$

due to (34). Thus, by Fact 2, the logarithmic derivative of f with respect to λ_r is

$$\begin{aligned} \frac{1}{f} \frac{\partial f}{\partial \lambda_r} &= \operatorname{trace}\bigl(H(\widetilde{A}) H(A_r) T(\widetilde{A})^{-1}\bigr) \\ &= \operatorname{trace}\bigl(H(A_r) T(\widetilde{A})^{-1} H(\widetilde{A})\bigr). \end{aligned} \tag{45}$$

In a similar way, one gets

$$\begin{aligned} \frac{1}{g} \frac{\partial g}{\partial \lambda_r} &= -\operatorname{trace}\bigl(T^{-1}\bigl(A^{-1}\bigr) H(A_r) H\bigl(\widetilde{A}^{-1}\bigr)\bigr) \\ &= -\operatorname{trace}\bigl(H(A_r) H\bigl(\widetilde{A}^{-1}\bigr) T^{-1}\bigl(A^{-1}\bigr)\bigr). \end{aligned} \tag{46}$$

From (35) it further follows that

$$0 = T(\widetilde{A})H\big(\widetilde{A}^{-1}\big) + H(\widetilde{A})T\big(A^{-1}\big)$$

and, consequently,

$$0 = H\big(\widetilde{A}^{-1}\big)T^{-1}\big(A^{-1}\big) + T^{-1}(\widetilde{A})H(\widetilde{A}).$$

In connection with Fact 1, this implies the equality of (45) and (46). Thus, the ratio of the functions f and g does not depend on λ_r. Since we already know that $f(\lambda_1, \ldots, \lambda_r) = g(\lambda_1, \ldots, \lambda_r)$ for $\lambda_r = 0$, it follows that $f(\lambda_1, \ldots, \lambda_r) = g(\lambda_1, \ldots, \lambda_r)$ for all complex parameters $\lambda_1, \ldots, \lambda_r$. $\square$

4. Banach algebras associated with suitable Banach algebras

4.1. Banach algebras of operators on $\ell^2(\mathbb{Z}^+)$

With every rigid and suitable Banach subalgebra $\mathcal{R}$ of $L(\ell^2(\mathbb{Z}))$, we associate a Banach algebra $\mathcal{O}(\mathcal{R})$ of operators on $\ell^2(\mathbb{Z}^+)$.

Theorem 4.1. *Let $\mathcal{R}$ be a rigid and suitable Banach subalgebra of $L(\ell^2(\mathbb{Z}))$. The set*

$$\mathcal{O}(\mathcal{R}) := \big\{T(A) + K\colon\ A \in \mathcal{R},\ K \in \mathcal{C}_1\big(\ell^2\big(\mathbb{Z}^+\big)\big)\big\}$$

is a unital Banach algebra with norm

$$\big\|T(A) + K\big\|_{\mathcal{O}(\mathcal{R})} := \|A\|_{\mathcal{R}} + \|K\|_{\mathcal{C}_1}.$$

Proof. The norm is correctly defined since $T(A) + K = 0$ implies $A = 0$ and $K = 0$ due to rigidity. Evidently, $\mathcal{O}(\mathcal{R})$ is a linear space. We are going to check that $\mathcal{Q}(\mathcal{R})$ is closed under multiplication. Let $A, B \in \mathcal{R}$ and $K, L \in \mathcal{C}_1$. Then using (34),

$$\begin{aligned}\big(T(A) + K\big)\big(T(B) + L\big) &= T(A)T(B) + T(A)L + KT(B) + KL \\ &= T(AB) + \big(T(A)L + KT(B) + KL - H(A)H(\widetilde{B})\big).\end{aligned}$$

This operator is in $\mathcal{O}(\mathcal{R})$ because the suitability of $\mathcal{R}$ implies that $H(A)H(\widetilde{B}) = PAQBP$ is trace class. Moreover, the norm can be estimated as follows:

$$\begin{aligned}&\big\|\big(T(A) + K\big)\big(T(B) + L\big)\big\|_{\mathcal{O}(\mathcal{R})} \\ &\quad = \|AB\|_{\mathcal{R}} + \big\|T(A)L + KT(B) + KL - PAQBP\big\|_{\mathcal{C}_1} \\ &\quad \leqslant 2\|A\|_{\mathcal{R}}\|B\|_{\mathcal{R}} + \|A\|_{L(\ell^2(\mathbb{Z}))}\|L\|_{\mathcal{C}_1} + \|B\|_{L(\ell^2(\mathbb{Z}))}\|K\|_{\mathcal{C}_1} + \|K\|_{\mathcal{C}_1}\|L\|_{\mathcal{C}_1} \\ &\quad \leqslant 2\|A\|_{\mathcal{R}}\|B\|_{\mathcal{R}} + \|A\|_{\mathcal{R}}\|L\|_{\mathcal{C}_1} + \|B\|_{\mathcal{R}}\|K\|_{\mathcal{C}_1} + \|K\|_{\mathcal{C}_1}\|L\|_{\mathcal{C}_1} \\ &\quad \leqslant 2\big(\|A\|_{\mathcal{R}} + \|K\|_{\mathcal{C}_1}\big)\big(\|B\|_{\mathcal{R}} + \|L\|_{\mathcal{C}_1}\big) \\ &\quad = 2\big\|T(A) + K\big\|_{\mathcal{O}(\mathcal{R})}\big\|T(B) + L\big\|_{\mathcal{O}(\mathcal{R})}.\end{aligned}$$

Finally, the completeness of $\mathcal{O}(\mathcal{R})$ follows from that of $\mathcal{R}$ and $\mathcal{C}_1(\ell^2(\mathbb{Z}^+))$ due to the definition of the norm on $\mathcal{O}(\mathcal{R})$. $\square$

For later purposes, let us make the following observation. Recall that given a Banach algebra $\mathcal{R}$ we have defined $\widetilde{\mathcal{R}}$ as the set of all operators $\widetilde{A}$ with $A \in \mathcal{R}$ with the norm $\|\widetilde{A}\|_{\widetilde{\mathcal{R}}} = \|A\|_{\mathcal{R}}$. The definition of rigidity and suitability guarantees that $\widetilde{\mathcal{R}}$ is suitable and rigid whenever so is $\mathcal{R}$. Hence we are able to consider $\mathcal{O}(\widetilde{\mathcal{R}})$ as well. Clearly,

$$\mathcal{O}(\widetilde{\mathcal{R}}) = \{T(\widetilde{A}) + K\colon\ A \in \mathcal{R},\ K \in \mathcal{C}_1(\ell^2(\mathbb{Z}^+))\} \tag{47}$$

with the norm

$$\|T(\widetilde{A}) + K\|_{\mathcal{O}(\widetilde{\mathcal{R}})} = \|A\|_{\mathcal{R}} + \|K\|_{\mathcal{C}_1}.$$

Both $\mathcal{O}(\mathcal{R})$ and $\mathcal{O}(\widetilde{\mathcal{R}})$ will be used later on.

4.2. *Banach algebras of sequences*

In this section, we associate with every rigid, suitable, and shift-invariant Banach algebra $\mathcal{R}$, and with each distinguished sequence h, a Banach algebras $\mathcal{S}_h(\mathcal{R})$ of sequences that arise from the finite sections of operators in $\mathcal{O}(\mathcal{R})$.

In the following theorem, we will need the reflection operators

$$W_n : \ell^2(\mathbb{Z}^+) \to \ell^2(\mathbb{Z}^+), \quad (x_0, x_1, \ldots) \mapsto (x_{n-1}, x_{n-2}, \ldots, x_0, 0, 0, \ldots).$$

Theorem 4.2. *Let $\mathcal{R}$ be a rigid, suitable, and shift-invariant Banach subalgebra of $L(\ell^2(\mathbb{Z}^+))$, and let $h \in \mathcal{H}$ be distinguished for $\mathcal{R}$. Then the set $\mathcal{S}_h(\mathcal{R})$ consisting of all sequences $(A_n)_{n=1}^{\infty}$ of operators $A_n : \operatorname{im} P_{h(n)} \to \operatorname{im} P_{h(n)}$ of the form*

$$A_n = P_{h(n)}T(A)P_{h(n)} + P_{h(n)}KP_{h(n)} + W_{h(n)}LW_{h(n)} + G_n \tag{48}$$

with $A \in \mathcal{R}_h$, $K, L \in \mathcal{C}_1(\ell^2(\mathbb{Z}^+))$, $G_n \in \mathcal{C}_1(\operatorname{im} P_{h(n)})$ and $\|G_n\|_{\mathcal{C}_1} \to 0$ forms a unital Banach algebra with respect to the operations

$$(A_n) + (B_n) := (A_n + B_n), \qquad (A_n)(B_n) := (A_nB_n), \qquad \lambda(A_n) := (\lambda A_n)$$

and the norm given by

$$\|(A_n)_{n\geqslant 1}\|_{\mathcal{S}_h(\mathcal{R})} := \|A\|_{\mathcal{R}} + \|K\|_{\mathcal{C}_1} + \|L\|_{\mathcal{C}_1} + \sup_{n\geqslant 1}\|G_n\|_{\mathcal{C}_1}. \tag{49}$$

Moreover, the set $\mathcal{J}_h(\mathcal{R})$ of all sequences (J_n) of the form

$$J_n = P_{h(n)}KP_{h(n)} + W_{h(n)}LW_{h(n)} + G_n$$

with $K, L \in \mathcal{C}_1(\ell^2(\mathbb{Z}^+))$, $G_n \in \mathcal{C}_1(\operatorname{im} P_{h(n)})$ and $\|G_n\|_{\mathcal{C}_1} \to 0$ forms a closed two-sided ideal of $\mathcal{S}_h(\mathcal{R})$.

Proof. First we show that the norm is correctly defined. Assume that

$$P_{h(n)}T(A)P_{h(n)} + P_{h(n)}KP_{h(n)} + W_{h(n)}LW_{h(n)} + G_n = 0$$

for all $n \in \mathbb{Z}^+$. Taking the strong limit $n \to \infty$ yields $T(A) + K = 0$. Because of rigidity, $A = 0$ and $K = 0$. Hence

$$W_{h(n)}LW_{h(n)} + G_n = 0$$

for all $n \in \mathbb{Z}^+$. Now multiply from both sides by $W_{h(n)}$, use $W_{h(n)}^2 = P_{h(n)}$, and take the strong limit again in order to obtain $L = 0$ and finally $G_n = 0$.

It is evident that $\mathcal{S}_h(\mathcal{R})$ is a linear space which is complete with respect to the defined norm. In order to check that it is an algebra consider

$$A_n = P_{h(n)}T(A)P_{h(n)} + P_{h(n)}K_1P_{h(n)} + W_{h(n)}L_1W_{h(n)} + G_n^{(1)},$$
$$B_n = P_{h(n)}T(B)P_{h(n)} + P_{h(n)}K_2P_{h(n)} + W_{h(n)}L_2W_{h(n)} + G_n^{(2)}.$$

We have to multiply each term in the first sum with each term in the second sum, and to show that the product is in $\mathcal{S}_h(\mathcal{R})$ again and moreover that the norm can be estimated appropriately. This is evident if one of the factors is $G_n^{(1)}$ or $G_n^{(2)}$. The other cases are considered in what follows. (For sake of brevity, we will omit some cases which are completely analogous to the ones considered.) First of all,

$$P_{h(n)}K_1P_{h(n)} \cdot P_{h(n)}K_2P_{h(n)} = P_{h(n)}K_1K_2P_{h(n)} - P_{h(n)}K_1Q_{h(n)}K_2P_{h(n)},$$
$$W_{h(n)}L_1W_{h(n)} \cdot W_{h(n)}L_2W_{h(n)} = W_{h(n)}L_1L_2W_{h(n)} - W_{h(n)}L_1Q_{h(n)}L_2W_{h(n)},$$
$$P_{h(n)}K_1P_{h(n)} \cdot W_{h(n)}L_2W_{h(n)} = P_{h(n)}K_1W_{h(n)}L_2W_{h(n)}.$$

Since $Q_{h(n)} := P - P_{h(n)}$ converges strongly to zero on $\ell^2(\mathbb{Z}^+)$ and $W_{h(n)}$ converges weakly to zero, the last terms in the first two equations as well as the term in the third equation converge to zero in the trace norm. Quite similarly,

$$P_{h(n)}T(A)P_{h(n)} \cdot P_{h(n)}K_2P_{h(n)} = P_{h(n)}T(A)K_2P_{h(n)} - P_{h(n)}T(A)Q_{h(n)}K_2P_{h(n)}$$

with the last term converging to zero.

The following two cases are slightly more complicated. Therein we are going to use the assumption that h is distinguished and that $\mathcal{R}$ is shift-invariant in order to establish that for $A \in \mathcal{R}$ we have $U_{-h(n)}AU_{h(n)} \in \mathcal{R}$ and

$$U_{-h(n)}AU_{h(n)} = A + C_n \tag{50}$$

where $C_n \to 0$ in the norm of $\mathcal{R}$. In particular, $C_n \to 0$ in the operator norm. We are also going to use the identities

$$W_{h(n)}P = P_{h(n)}PJU_{-h(n)}, \qquad PW_{h(n)} = U_{h(n)}JPP_{h(n)}. \tag{51}$$

First we consider the product

$$\begin{aligned}
&\big(P_{h(n)}T(A)P_{h(n)}\big)\cdot\big(W_{h(n)}L_2W_{h(n)}\big)\\
&\quad= W_{h(n)}(W_{h(n)}PAPW_{h(n)})L_2W_{h(n)}\\
&\quad= W_{h(n)}\big(P_{h(n)}PJ(U_{-h(n)}AU_{h(n)})JPP_{h(n)}\big)L_2W_{h(n)}\\
&\quad= W_{h(n)}(PJAJPL_2)W_{h(n)}+G_n'\\
&\quad= W_{h(n)}T(\widetilde{A})L_2W_{h(n)}+G_n'
\end{aligned}$$

with $T(\widetilde{A})L_2\in\mathcal{C}_1(\ell^2(\mathbb{Z}^+))$ and

$$\begin{aligned}
G_n' = &-W_{h(n)}PJU_{-h(n)}AU_{h(n)}JPQ_{h(n)}L_2W_{h(n)}\\
&+W_{h(n)}PJ(U_{-h(n)}AU_{h(n)}-A)JPL_2W_{h(n)}.
\end{aligned}$$

The sequence G_n tends to zero in the trace norm because $Q_{h(n)}\to 0$ strongly, $U_{-h(n)}AU_{h(n)}\to A$ in the operator norm, and L_2 is trace class. A careful examination, using in particular $\|U_{-h(n)}AU_{h(n)}\|_{\mathcal{R}}=\|A\|_{\mathcal{R}}$, shows that we can estimate

$$\|G_n'\|_{\mathcal{C}_1}\leqslant 3\|A\|_{\mathcal{R}}\|L_2\|_{\mathcal{C}_1}.$$

Let us finally consider

$$\begin{aligned}
&\big(P_{h(n)}T(A)P_{h(n)}\big)\cdot\big(P_{h(n)}T(B)P_{h(n)}\big)\\
&\quad= P_{h(n)}T(AB)P_{h(n)}-P_{h(n)}H(A)H(\widetilde{B})P_{h(n)}-P_{h(n)}T(A)Q_{h(n)}T(B)P_{h(n)}.
\end{aligned}$$

Here we used (34). The term $H(A)H(\widetilde{B})$ is trace class because of the assumption that $\mathcal{R}$ is suitable. Using (51) and $Q_{h(n)}=U_{h(n)}PU_{-h(n)}$, the last term in the sum can be rewritten as

$$\begin{aligned}
&W_{h(n)}W_{h(n)}PAPQ_{h(n)}PBPW_{h(n)}W_{h(n)}\\
&\quad= W_{h(n)}JQU_{-h(n)}AU_{h(n)}PU_{-h(n)}BU_{h(n)}QJW_{h(n)}.
\end{aligned}\tag{52}$$

Now we employ the observation made in connection with (50) and substitute $U_{-h(n)}AU_{h(n)}=A+C_n^{(1)}$ and $U_{-h(n)}BU_{h(n)}=B+C_n^{(2)}$, where all terms belong to $\mathcal{R}$ and the $C_n^{(i)}$ converge to zero in the norm of $\mathcal{R}$. Hence (52) equals

$$W_{h(n)}H\big(\widetilde{A+C_n^{(1)}}\big)H\big(B+C_n^{(2)}\big)W_{h(n)},$$

which is $W_{h(n)}H(\widetilde{A})H(B)W_{h(n)}$ plus three terms whose trace norm can be estimated (because of property (ii) of suitability) by a constant times

$$\|A\|_{\mathcal{R}}\cdot\big\|C_n^{(2)}\big\|_{\mathcal{R}}+\big\|C_n^{(1)}\big\|_{\mathcal{R}}\cdot\|B\|_{\mathcal{R}}+\big\|C_n^{(1)}\big\|_{\mathcal{R}}\cdot\big\|C_n^{(2)}\big\|_{\mathcal{R}}.$$

These three terms converge to zero in the trace norm, and $H(\widetilde{A})H(B)$ is trace class. In summarizing the previous steps we obtain

$$\big(P_{h(n)}T(A)P_{h(n)}\big)\cdot\big(P_{h(n)}T(B)P_{h(n)}\big) = P_{h(n)}T(AB)P_{h(n)} - P_{h(n)}H(A)H(\widetilde{B})P_{h(n)} - W_{h(n)}H(\widetilde{A})H(B)W_{h(n)} + G_n'' \tag{53}$$

with $\|G_n''\|_{\mathcal{C}_1} \to 0$.

In conclusion, the product of the above defined sequences (A_n) and (B_n) can be written as

$$A_nB_n = P_{h(n)}T(AB)P_{h(n)} + P_{h(n)}KP_{h(n)} + W_{h(n)}LW_{h(n)} + G_n \tag{54}$$

with

$$K = K_1K_2 + T(A)K_2 + K_1T(B) - H(A)H(\widetilde{B}), \tag{55}$$

$$L = L_1L_2 + T(\widetilde{A})L_2 + L_1T(\widetilde{B}) - H(\widetilde{A})H(B) \tag{56}$$

and where G_n is a sequence of trace class operators tending to zero in the trace norm. Elaborating on the precise expression for G_n one can actually show that

$$\big\|(A_n)(B_n)\big\|_{\mathcal{S}_h(\mathcal{R})} \leqslant M\cdot\big\|(A_n)\big\|_{\mathcal{S}_h(\mathcal{R})}\big\|(B_n)\big\|_{\mathcal{S}_h(\mathcal{R})}$$

with some constant M. This concludes the proof that $\mathcal{S}_h(\mathcal{R})$ is a Banach algebra.

The fact that $\mathcal{J}_h(\mathcal{R})$ is closed follows immediately from the definition of the norm and that it is an ideal follows from the formulas (54)–(56). □

Theorem 4.3. *Under the assumptions of the previous theorem, the mappings $\mathcal{W}_h$ and $\widetilde{\mathcal{W}}_h$ defined by*

$$\mathcal{W}_h : (A_n)_{n\geqslant 1} \in \mathcal{S}_h(\mathcal{R}) \mapsto T(A) + K \in \mathcal{O}(\mathcal{R}), \tag{57}$$

$$\widetilde{\mathcal{W}}_h : (A_n)_{n\geqslant 1} \in \mathcal{S}_h(\mathcal{R}) \mapsto T(\widetilde{A}) + L \in \mathcal{O}(\widetilde{\mathcal{R}}), \tag{58}$$

where the sequences (A_n) of the form (48)*, are unital Banach algebra homomorphisms.*

Proof. In view of the definition of the norms by (49) and

$$\big\|T(A)+K\big\|_{\mathcal{O}(\mathcal{R})} = \|A\|_{\mathcal{R}} + \|K\|_{\mathcal{C}_1}, \qquad \big\|T(\widetilde{B})+L\big\|_{\mathcal{O}(\widetilde{\mathcal{R}})} = \|B\|_{\mathcal{R}} + \|L\|_{\mathcal{C}_1},$$

the continuity of the mappings is obvious. The linearity is also clear. Their multiplicativity follows from formulas (54)–(56) in connection with (34). □

5. The Strong Szegö–Widom Limit Theorem

Before stating our main result, the generalization of the Strong Szegö–Widom Limit Theorem, we need an auxiliary fact, namely, a generalization of Proposition 9.2 from [12]. It is in some way the most important part of the Banach algebra approach, where its main idea is exhibited.

Proposition 5.1. *Let $\mathcal{R}$ be a rigid, suitable, and shift-invariant Banach algebra, let $h \in \mathcal{H}$, and let $A_1, \dots, A_r \in \mathcal{R}$. Then the sequence $(B_n)_{n\geqslant 1}$ defined by*

$$B_n := P_{h(n)} T\big(e^{A_1} \cdots e^{A_r}\big) P_{h(n)} \cdot e^{-P_{h(n)}T(A_r)P_{h(n)}} \cdots e^{-P_{h(n)}T(A_1)P_{h(n)}} \tag{59}$$

belongs to $\mathcal{S}_h(\mathcal{R})$. Moreover, there exist operators $K, L \in \mathcal{C}_1(\ell^2(\mathbb{Z}^+))$ and $G_n \in \mathcal{C}_1(\operatorname{im} P_{h(n)})$ with $\|G_n\|_{\mathcal{C}_1} \to 0$ such that

$$B_n = P_{h(n)} + P_{h(n)} K P_{h(n)} + W_{h(n)} L W_{h(n)} + G_n. \tag{60}$$

The operators K and L are determined by

$$P + K = T\big(e^{A_1} \cdots e^{A_r}\big) \cdot e^{-T(A_r)} \cdots e^{-T(A_1)},$$
$$P + L = T\big(e^{\widetilde{A_1}} \cdots e^{\widetilde{A_r}}\big) \cdot e^{-T(\widetilde{A_r})} \cdots e^{-T(\widetilde{A_1})}.$$

Proof. We first observe the general fact that if $(A_n) \in \mathcal{S}_h(\mathcal{R})$, then $(e^{A_n}) = e^{(A_n)}$. From this it follows that the sequence (B_n) belongs to $\mathcal{S}_h(\mathcal{R})$ and that

$$(B_n)_{n\geqslant 1} = \big(P_{h(n)} T\big(e^{A_1} \cdots e^{A_r}\big) P_{h(n)}\big)_{n\geqslant 1} \cdot e^{-(P_{h(n)}T(A_r)P_{h(n)})_{n\geqslant 1}} \cdots e^{-(P_{h(n)}T(A_1)P_{h(n)})_{n\geqslant 1}}.$$

We can write this sequence as

$$(B_n)_{n\geqslant 1} = \Lambda\big(e^{A_1} \cdots e^{A_r}\big) \cdot e^{-\Lambda(A_r)} \cdots e^{-\Lambda(A_1)}, \tag{61}$$

where Λ is the mapping

$$\mathcal{R} \to \mathcal{S}_h(\mathcal{R}), \qquad A \mapsto \big(P_{h(n)} T(A) P_{h(n)}\big)_{n\geqslant 1}.$$

Evidently, Λ is linear and bounded. Let Φ denote the canonical homomorphism

$$\Phi : \mathcal{S}_h(\mathcal{R}) \to \mathcal{S}_h(\mathcal{R})/\mathcal{J}_h(\mathcal{R}), \qquad (A_n) \mapsto (A_n) + \mathcal{J}_h(\mathcal{R}).$$

It follows from (53) that the mapping

$$\Phi \circ \Lambda : \mathcal{R} \to \mathcal{S}_h(\mathcal{R})/\mathcal{J}_h(\mathcal{R})$$

is a continuous homomorphism between Banach algebras. Applying Φ to both sides of (61) yields

$$\begin{aligned}
\Phi\big((B_n)_{n\geqslant 1}\big) &= (\Phi \circ \Lambda)\big(e^{A_1} \cdots e^{A_r}\big) \cdot e^{-(\Phi\circ\Lambda)(A_r)} \cdots e^{-(\Phi\circ\Lambda)(A_1)} \\
&= (\Phi \circ \Lambda)\big(e^{A_1} \cdots e^{A_r} e^{-A_r} \cdots e^{-A_1}\big) \\
&= (\Phi \circ \Lambda)(I) \\
&= (P_{h(n)})_{n\geqslant 1} + \mathcal{J}_h(\mathcal{R}).
\end{aligned}$$

Hence, $(B_n - P_{h(n)})_{n\geqslant 1}$ is in $\mathcal{J}_h$, which proves (60). The representations of $I+K$ and $I+L$ follow by applying the homomorphisms $\mathcal{W}_h$ and $\widetilde{\mathcal{W}}_h$ (which were defined in Theorem 4.3) to both sides of

$$(B_n)_{n\geqslant 1} = (P_{h(n)} + P_{h(n)}KP_{h(n)} + W_{h(n)}LW_{h(n)} + G_n)_{n\geqslant 1},$$

respectively. □

A consequence of the previous proposition is the theorem that we will state next. It is the first part of our main result. Therein the asymptotics of determinants are reduced to the asymptotics of traces. The asymptotics of the traces is left unevaluated, which allows us to stated the result in greater generality.

For any Banach algebra $\mathcal{R}$ we denote by $\mathcal{G}_1\mathcal{R}$ the connected component of the group of all invertible elements in $\mathcal{R}$ which contains the unit element. It is a well-known result of Lorch – see, e.g., Theorem 3.3.7 in [1], or [20] – that $\mathcal{G}_1\mathcal{R}$ is the set of all finite products of exponentials of elements in $\mathcal{R}$. Hence the basic assumption on the symbol A is that it belongs to $\mathcal{G}_1\mathcal{R}$. This corresponds precisely to the basic assumption in the classical Szegö–Widom Theorem (see Theorem 1.1 and the remarks afterwards).

There is, however, a subtlety in the following statement. The operators $A_1, \dots, A_r$ are not uniquely determined by A, and they do appear in the limit expression.

Theorem 5.2. *Let $\mathcal{R}$ be a rigid, suitable, and shift-invariant Banach algebra, and let $h \in \mathcal{H}$ be a distinguished sequence for $\mathcal{R}$. If $A_1, \dots, A_r \in \mathcal{R}$, and*

$$A = e^{A_1} \cdots e^{A_r},$$

then

$$\lim_{n\to\infty} \frac{\det(P_{h(n)}T(A)P_{h(n)})}{\exp(\operatorname{trace}(P_{h(n)}T(A_1 + \cdots + A_r)P_{h(n)}))} = \det T(A)T\big(A^{-1}\big).$$

Proof. Consider the sequence (B_n) defined by (59). Take the determinant and observe that

$$\begin{aligned}\det B_n &= \det\big(P_{h(n)}T(A)P_{h(n)} \cdot e^{-P_{h(n)}T(A_r)P_{h(n)}} \cdots e^{-P_{h(n)}T(A_1)P_{h(n)}}\big) \\ &= \det\big(P_{h(n)}T(A)P_{h(n)}\big) \cdot e^{-\operatorname{trace}(P_{h(n)}T(A_r)P_{h(n)})} \cdots e^{-\operatorname{trace}(P_{h(n)}T(A_1)P_{h(n)})} \\ &= \det\big(P_{h(n)}T(A)P_{h(n)}\big) \cdot e^{-\operatorname{trace}(P_{h(n)}T(A_r+\cdots+A_1)P_{h(n)})}.\end{aligned}$$

By Proposition 5.1, the sequence B_n is of the form (60). A basic observation (see, e.g., Lemmas 9.1 and 9.3 of [12]) implies

$$\lim_{n\to\infty} \det B_n = \det(P+K)\cdot\det(P+L). \tag{62}$$

Again from Proposition 5.1 we infer that the right-hand side is equal to the product of two operator determinants,

$$\det T(A)\cdot e^{-T(A_1)}\cdots e^{-T(A_r)}$$

and

$$\det T(\widetilde{A}) \cdot e^{-T(\widetilde{A_1})} \cdots e^{-T(\widetilde{A_r})}.$$

The last determinant can be rewritten, using Theorem 3.8, as

$$\det T\bigl(A^{-1}\bigr) \cdot e^{T(A_r)} \cdots e^{T(A_1)}.$$

Thus both determinants multiplied together yield the constant $\det T(A)T(A^{-1})$. Remark that $T(A)T(A^{-1}) = I - H(A)H(\widetilde{A^{-1}})$ with the product of the generalized Hankel operators being trace class. □

The second part of our main result is the following. Recall that $M(a)$ stands for the mean of an almost periodic function, and $D(B)$ stands for the main diagonal of an operator $B \in L(\ell^2(\mathbb{Z}))$.

Theorem 5.3 *(Generalized Strong Szegö–Widom Limit Theorem). Let $\boldsymbol{\beta}$ be an admissible and compatible weight on an additive subgroup Ξ of $\mathbb{R}/\mathbb{Z}$. Let $\mathcal{R} = \mathcal{W}_{\alpha_1,\alpha_2}(APW(\mathbb{Z}, \Xi, \boldsymbol{\beta}))$ with $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. Suppose that $h \in \mathcal{H}$ is a distinguished sequence for Ξ. If $A_1, \ldots, A_r \in \mathcal{R}$, and*

$$A = e^{A_1} \cdots e^{A_r}, \tag{63}$$

then

$$\lim_{n\to\infty} \frac{\det(P_{h(n)}T(A)P_{h(n)})}{G^{h(n)}} = \det T(A)T\bigl(A^{-1}\bigr), \tag{64}$$

with

$$G := \exp M(a), \quad a := D(A_1 + \cdots + A_r). \tag{65}$$

Proof. Let us first remark that due to Theorem 3.5 the Banach algebra is rigid, suitable, and shift-invariant. Moreover, h is distinguished for $\mathcal{R}$ (see Propositions 2.1 and 3.6, and Theorem 2.6). Hence Theorem 5.2 can be applied. We are left with the asymptotics of

$$\operatorname{trace}\bigl(P_{h(n)}T(A_1 + \cdots + A_r)P_{h(n)}\bigr) = \sum_{k=0}^{h(n)-1} a(k)$$

with $a \in APW(\mathbb{Z}, \Xi, \boldsymbol{\beta})$ being defined as above. Now it remains to apply Theorem 2.7. □

We conclude this section with a couple of observations and questions.

Ambiguity of the constant G. The constant G is defined only implicitly in terms of A. More precisely, different choices for $A_1, \ldots, A_r$ could yield the same A, but different constants G. As the following example shows this can happen in the periodic case. We conjecture that this can always happen in the almost periodic case.

For the periodic case (with period $N=2$) we can give the following simple example ($r=1$). Choose $A=e^{A_1}=e^{A_2}=I$ with $A_1=0$, $A_2=aI$,

$$a(n)=\begin{cases}2\pi i & n \text{ even},\\ 0 & n \text{ odd},\end{cases}$$

i.e., $a=\pi i\boldsymbol{e}_0+\pi i\boldsymbol{e}_{1/2}$. The corresponding constants $G_1=1$ and $G_2=-1$ are different.

On the other hand, since the constant G describes an asymptotics, it cannot be completely arbitrary. The modification which seem to be "admissible" is replacing G by $e^{2\pi i\xi}\cdot G$ for arbitrary $\xi\in\Xi$. In this connection recall that h being distinguished implies that $e^{2\pi i\xi h(n)}\to 1$ so that in fact there is no change in the asymptotic description (64).

Problem of inverse closedness. It seems practically very difficult to decide whether an operator belongs to $\mathcal{G}_1\mathcal{R}$ with $\mathcal{R}=\mathcal{W}_{\alpha_1,\alpha_2}(APW(\mathbb{Z},\Xi,\boldsymbol{\beta}))$. The question arises whether (under reasonable assumptions) the Banach algebra $\mathcal{R}$ is inverse closed in $L(\ell^2(\mathbb{Z}))$, i.e.,

$$\mathcal{G}\mathcal{R}=\mathcal{R}\cap\mathcal{G}\bigl(L\bigl(\ell^2(\mathbb{Z})\bigr)\bigr).$$

If this would be true, then it is conceivable that

$$\mathcal{G}_1\mathcal{R}=\mathcal{R}\cap\mathcal{G}_1\bigl(L\bigl(\ell^2(\mathbb{Z})\bigr)\bigr).$$

One can then replace the assumption by a somewhat more tractable one. For instance, if $A\in\mathcal{R}$ is self-adjoint and its spectrum $\mathrm{sp}(A)$ (in $L(\ell^2(\mathbb{Z}^+))$) is known, then the theorem applies to $A-\lambda I$ whenever $\lambda\notin\mathrm{sp}(A)$.

Of course, the almost Mathieu operator, which serves as the main example, is self-adjoint.

Let us remark that a necessary condition for the inverse closedness of $\mathcal{R}$ in $L(\ell^2(\mathbb{Z}))$ is the inverse closedness of $APW(\mathbb{Z},\Xi,\boldsymbol{\beta})$ in $AP(\mathbb{Z})$ (hence in $\ell^\infty(\mathbb{Z})$). In the case of finitely generated groups Ξ, this seems to be the case if the weight $\boldsymbol{\beta}$ does not grow exponentially.

6. Counterexamples to the asymptotics of the traces

The goal of this section is to show that the asymptotics

$$\sum_{k=0}^{h(n)-1}a(k)=h(n)\cdot M(a)+o(1),\quad n\to\infty, \tag{66}$$

does *not* hold in general, i.e., for general $a\in AP(\mathbb{Z})$ and $h:\mathbb{Z}^+\to\mathbb{Z}^+$ assumed to be distinguished (say, for the group generated by the Fourier spectrum of a).

As has been pointed out in Section 2, the asymptotics (66) does hold if the Fourier spectrum of a is finite, i.e., if the sum (16) is finite. It also holds "generically" if the Fourier spectrum of a belongs, for instance, to a singly-generated group $\Xi_\xi:=(\xi\mathbb{Z})/\mathbb{Z}$, assuming that the Fourier coefficients decay sufficiently fast (Theorems 2.7 and 2.8). More precisely, for almost every $\xi\in\mathbb{R}\setminus\mathbb{Q}$, there exists $\omega>0$ such that the asymptotics (66) holds for every

$$a=\sum_{k=-\infty}^{\infty}a_k\boldsymbol{e}_{k\xi}\quad\text{satisfying}\quad\sum_{k=-\infty}^{\infty}|a_k|\bigl(1+|k|\bigr)^{\omega}<\infty. \tag{67}$$

In contrast to that, we are going to show in what follows that there exists $\xi \in \mathbb{R} \setminus \mathbb{Q}$ and a distinguished sequence for Ξ_ξ such that the asymptotics (66) does *not* hold for certain a of the form (67). The sequences a we are going to consider have exponentially decaying Fourier coefficients. Moreover, we assume $a_k = 0$ for $k \leqslant 0$, i.e.,

$$a = \sum_{k=1}^{\infty} a_k \boldsymbol{e}_{k\xi}.$$

Notice that $M(a) = a_0 = 0$.

From the definition of the mean it follows that the error term in (66) is always $o(h(n))$. Our results suggest that not much improvement is possible in general.

We start with constructing irrational numbers ξ which will provide the basis for our counterexamples. They depend on a parameter $c > 1$, which we can later relate to the exponential decay of the Fourier coefficients a_k.

In what follows we denote by $\{x\}$ the fractional part of a real number x.

Proposition 6.1. *Given $c > 1$, there exist $\xi \in \mathbb{R}$ and strictly increasing sequences $\{q_d\}_{d=1}^{\infty}$ and $\{w_d\}_{d=1}^{\infty}$ of natural numbers such that*

$$w_d = q_1 \cdots q_d, \qquad \lim_{d\to\infty} \frac{q_{d+1}}{c^{w_d}} = 1, \qquad \xi = \sum_{d=1}^{\infty} \frac{1}{w_d}, \tag{68}$$

$$\{w_d \xi\} = \frac{1}{q_{d+1}}\left(1 + O\left(\frac{1}{q_{d+1}}\right)\right), \quad d \to \infty, \tag{69}$$

and

$$\|k\xi\|_{\mathbb{R}/\mathbb{Z}} \geqslant \frac{1}{w_d} - k \cdot O\left(\frac{1}{w_{d+1}}\right), \tag{70}$$

whenever $w_d \nmid k$.

Proof. Starting with any q_1 and $p_1 = 1$ we define recursively

$$w_d = q_1 \cdots q_d, \qquad q_{d+1} = \left[c^{w_d}\right] + s_d, \qquad p_{d+1} = p_d q_{d+1} + 1,$$

where $[x]$ denotes the integer part of $x \in \mathbb{R}$ and $s_d \in \mathbb{Z}^+$ is chosen such that

(i) w_d and p_d are co-prime,
(ii) $w_d < s_d \leqslant 2w_d$

for each $d \in \mathbb{N}$. Clearly, (ii) will guarantee that q_d and w_d are strictly increasing and that $q_{d+1} \sim c^{w_d}$ as $d \to \infty$.

To show that such a choice of s_d is possible we will argue by induction. The case $d = 1$ is obvious. Assume (i) holds for d. Then, as we will see shortly, there exists s_d satisfying (ii) such that (i) holds for $d + 1$. Indeed, when trying to find q_{d+1}, let us require $p_d q_{d+1} \not\equiv -1$ modulo each prime factor of w_d. Because $\gcd(w_d, p_d) = 1$ these conditions are equivalent to a system of congruences for q_{d+1} (or, equivalently, for s_d) modulo each prime factor of w_d. By the Chinese Remainder Theorem such a system of congruences has a unique solution modulo the product of the prime factors, and thus we can find the desired s_d and q_{d+1}. It follows that $p_{d+1} = p_d q_{d+1} + 1 \not\equiv 0$ modulo each prime factor of w_d, which implies that $\gcd(p_{d+1}, w_d) = 1$. Clearly, also $\gcd(p_{d+1}, q_{d+1}) = 1$. Thus we obtain condition (i) for $d+1$.

Let us prove (69) and (70). The recursion for p_d can be rewritten as

$$\frac{p_{d+1}}{w_{d+1}} = \frac{p_d}{w_d} + \frac{1}{w_{d+1}}.$$

Hence

$$\sum_{j=1}^{d} \frac{1}{w_j} = \frac{p_d}{w_d}, \qquad \xi = \frac{p_d}{w_d} + \sum_{j=d+1}^{\infty} \frac{1}{w_j}.$$

Multiplying ξ with w_d equals p_d plus the error term

$$\sum_{j=d+1}^{\infty} \frac{w_d}{w_{j+1}} = \frac{1}{q_{d+1}} + \frac{1}{q_{d+1}q_{d+2}} + \frac{1}{q_{d+1}q_{d+2}q_{d+3}} + \cdots.$$

Since q_d is increasing this series can be estimated by a geometric series, and we obtain (69).

Now observe that $\|kp_d/w_d\|_{\mathbb{R}/\mathbb{Z}} \geqslant 1/w_d$ if k is not a multiple of w_d (see also (i)). Moreover, using the same estimate as above,

$$\sum_{j=d+1}^{\infty} \frac{1}{w_j} = O\left(\frac{1}{w_{d+1}}\right).$$

This proves (70). $\square$

The properties stated in (68) and (69) imply that

$$\|w_d \xi\|_{\mathbb{R}/\mathbb{Z}} = O\left(q_{d+1}^{-1}\right) = O\left(c^{-w_d}\right), \quad d \to \infty,$$

which means that ξ is a Liouville number (hence it is irrational). Moreover, if d_n is an increasing sequence and k_n is some sequence such that $h(n) := k_n w_{d_n}$ is strictly increasing and

$$k_n = o(q_{d_n+1}) \quad \left(\text{or, equivalently, } k_n = o\left(c^{w_{d_n}}\right)\right) \text{ as } n \to \infty, \tag{71}$$

then, using property (69), we have

$$e^{2\pi i \xi h(n)} \to 1, \quad \text{as } n \to \infty. \tag{72}$$

Hence $h(n)$ defined in such a way is a distinguished sequence for $\Xi_\xi = (\xi\mathbb{Z})/\mathbb{Z}$.

Proposition 6.2. *Let $c > 1$, $\xi \in \mathbb{R}$, $\{q_d\}_{d=1}^{\infty}$ and $\{w_d\}_{d=1}^{\infty}$, be as in the previous proposition. Let $b > 1$ and $m \in \mathbb{N}$ be such that $c \leqslant b^{2m}$. Then*

$$\sum_{k\in\mathbb{N}\setminus S_m} \frac{b^{-k}}{|1-e^{2\pi i\xi k}|} < \infty, \tag{73}$$

where $S_m = \{kw_d\colon\ 1 \leqslant k \leqslant 2m,\ d \in \mathbb{N}\}$.

Proof. For sufficiently large d, say $d \geqslant d_0$, we have $(2m+1)w_{d-1} < w_d$ and $q_{d+1} \geqslant 2\gamma(2m+1)w_d$, where γ is the constant implied in (70). Using this last estimate and (70) we conclude that

$$\|k\xi\|_{\mathbb{R}/\mathbb{Z}} \geqslant \frac{1}{w_d} - \gamma \cdot \frac{(2m+1)w_d}{w_{d+1}} \geqslant \frac{1}{2w_d} \tag{$*$}$$

whenever $d \geqslant d_0$ and $k \in \mathbb{N}$ is such that $w_d \nmid k$ and $k \leqslant (2m+1)w_d$. In regard to the sum (73) we can omit finitely many terms and decompose the set of the remaining indices into

$$\bigl\{(2m+1)w_{d_0-1}, \ldots\bigr\} \setminus S_m = \bigcup_{d=d_0}^{\infty} A_d \cup B_d$$

with

$$\begin{aligned} A_d &= \bigl\{(2m+1)w_{d-1}, \ldots, w_d - 1\bigr\}, \\ B_d &= \bigl\{w_d, \ldots, (2m+1)w_d - 1\bigr\} \setminus \{kw_d\colon\ 1 \leqslant k \leqslant 2m\}. \end{aligned}$$

For $k \in A_d \cup B_d$ we have $|1-e^{2\pi i\xi k}|^{-1} = O(w_d)$; see $(*)$. Restricting the sum (73) to all $k \in A_d$ and all $d \geqslant d_0$ yields an upper estimate of a constant times

$$\begin{aligned} \sum_{d=d_0}^{\infty} b^{-(2m+1)w_{d-1}}(w_d)^2 &\leqslant C \sum_{d=d_0}^{\infty} b^{-(2m+1)w_{d-1}} c^{2w_{d-1}}(w_{d-1})^2 \\ &\leqslant C \sum_{d=d_0}^{\infty} b^{-w_{d-1}}(w_{d-1})^2 < \infty. \end{aligned}$$

Restricting the sum (73) to all $k \in B_d$ and all $d \geqslant d_0$ yields an upper estimate of a constant times

$$2m \sum_{d=d_0}^{\infty} b^{-w_d}(w_d)^2 < \infty.$$

Thus (73) is finite. This concludes the proof. $\square$

In view of the following theorem, remark that the assumptions therein ensure that h is a distinguished sequence.

Furthermore, this theorem has to be understood as follows. For given $c > 1$ we construct an irrational number ξ as in Proposition 6.1. Then we consider particular distinguished sequences, namely those characterized by (74). The class of such distinguished sequences h is still quite large. Independent of h, we consider almost periodic sequences whose Fourier coefficients behave roughly as b^{-k} where $b \in (1, c)$. More precisely, (75) and (76) hold. Then we are able to predict the asymptotics. If this asymptotics is different from $o(1)$, then we have found a counterexample to (66).

Theorem 6.3. *Let $c > 1$, $\xi \in \mathbb{R}$, $\{q_d\}_{d=1}^{\infty}$ and $\{w_d\}_{d=1}^{\infty}$ be as in Proposition* 6.1. *Let $h(n) = k_n w_{d_n}$ be strictly increasing with $d_n, k_n \in \mathbb{N}$ such that $d_n \to \infty$ and*

$$\frac{k_n w_{d_n}}{q_{d_n+1}} = o(1), \quad n \to \infty. \tag{74}$$

Moreover, let

$$a = \sum_{k \geqslant 1} a_k \boldsymbol{e}_{k\xi}, \tag{75}$$

and assume $1 < b < c$ such that

$$b = \liminf_{k \to \infty} |a_k|^{-1/k} = \lim_{n \to \infty} |a_{w_{d_n}}|^{-1/w_{d_n}}. \tag{76}$$

Then

$$\sum_{j=0}^{h(n)-1} a(j) = a_{w_{d_n}} w_{d_n} k_n \bigl(1 + o(1)\bigr) + o(1), \quad n \to \infty. \tag{77}$$

Proof. First of all observe that

$$\sum_{j=0}^{h(n)-1} a(j) = \sum_{k=1}^{\infty} a_k \frac{1 - e^{2\pi i \xi k h(n)}}{1 - e^{2\pi i \xi k}}. \tag{78}$$

Each term in this sum converges to zero because $e^{2\pi i \xi h(n)} \to 1$ as $n \to \infty$. Clearly, (74) implies (71), and hence $h(n)$ is a distinguished sequence.

Let $0 < \varepsilon < 1/3$ be such that $b^{1+\varepsilon} < c$, and choose $m \in \mathbb{N}$ such that $c \leqslant b^{(2m+1)(1-\varepsilon)}$. Because of (76) we have

$$b^{-k(1+\varepsilon)} \leqslant |a_k| \leqslant b^{-k(1-\varepsilon)}$$

for all sufficiently large k, say $k \geqslant K_0$. We split the above sum (78) into a finite sum over k (which is clearly $o(1)$ as $n \to \infty$) and

$$\sum_{d=d_0}^{\infty} \sum_{k=1}^{2m} a_{kw_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} + \sum_{\substack{k \in \mathbb{N} \setminus S_m \\ k \geqslant K_0}} a_k \frac{1 - e^{2\pi i \xi k h(n)}}{1 - e^{2\pi i \xi k}}$$

with d_0 sufficiently large and fixed. In fact, we choose d_0 such that

(a) the error term $O(1/q_{d+1})$ in (69) is bounded by $1/2$ for all $d \geqslant d_0$,
(b) $q_{d+1} > 3m$ for all $d \geqslant d_0$,
(c) $w_{d_0} \geqslant K_0$.

Using Proposition 6.2 with $b^{1-\varepsilon}$ instead of b, it follows that the second term converges to zero as $n \to \infty$ because of dominated convergence. Thus we obtain

$$\sum_{d=d_0}^{\infty} \sum_{k=1}^{2m} a_{kw_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} + o(1), \quad n \to \infty.$$

Now assume that n is large enough to ensure that $d_n \geqslant d_0$. Then one part of the above sum can be estimated as follows,

$$\begin{aligned}
\left| \sum_{d=d_n+1}^{\infty} \sum_{k=1}^{2m} a_{kw_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} \right| &\leqslant \sum_{d=d_n+1}^{\infty} \sum_{k=1}^{2m} h(n) |a_{kw_d}| \\
&\leqslant 2mh(n) \sum_{d=d_n+1}^{\infty} b^{-w_d(1-\varepsilon)} \\
&\leqslant 2mh(n) \frac{b^{-w_{d_n+1}(1-\varepsilon)}}{1 - b^{-(1-\varepsilon)}} \\
&= O\big(w_{d_n+1} b^{-w_{d_n+1}(1-\varepsilon)}\big) = o(1), \quad n \to \infty.
\end{aligned}$$

Here we have used $h(n) = o(q_{d_n+1}) = O(w_{d_n+1})$. Thus we are left with

$$\sum_{d=d_0}^{d_n} \sum_{k=1}^{2m} a_{kw_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} + o(1), \quad n \to \infty.$$

Henceforth we will assume $d_0 \leqslant d \leqslant d_n$ and $1 \leqslant k \leqslant 2m$. Because of Proposition 6.1, formulas (69) and (70), we have the estimates

$$\{\xi k w_d\} = \frac{k}{q_{d+1}} \left(1 + O\left(\frac{1}{q_{d+1}}\right)\right)$$

and

$$\big\{\xi k w_d h(n)\big\} = \{\xi k w_d k_n w_{d_n}\} = \frac{w_d k k_n}{q_{d_n+1}}\left(1 + O\left(\frac{1}{q_{d_n+1}}\right)\right)$$

with the error term bounded by $1/2$ because of (a). For the first estimate, we have used that $k/q_{d+1} \leqslant 2m/q_{d+1} < 2/3$, see (b). For the second estimate we have used that $w_d k k_n/q_{d_n+1} \leqslant 2m w_{d_n} k_n/q_{d_n+1} = o(1)$, by (74), which is less than $2/3$ for n sufficiently large. We can conclude that

$$\frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} = \frac{q_{d+1} w_d k_n}{q_{d_n+1}}\left(1 + O\left(\frac{1}{q_{d+1}}\right)\right) = O\left(\frac{q_{d+1} w_d k_n}{q_{d_n+1}}\right).$$

This bound is uniform in k, d, and d_n, as long as n is sufficiently large.

In the special case $k = 1$ and $d = d_n$, we get

$$a_{w_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} = a_{w_{d_n}} k_n w_{d_n}\left(1 + O\left(\frac{1}{q_{d_n+1}}\right)\right),$$

which yields the leading term. In general, we have to sum over the terms

$$a_{k w_d} \frac{1 - e^{2\pi i \xi k w_d h(n)}}{1 - e^{2\pi i \xi k w_d}} = O\left(\frac{a_{k w_d} q_{d+1} w_d k_n}{q_{d_n+1}}\right) = a_{w_{d_n}} k_n w_{d_n} O\left(\frac{a_{k w_d} w_d q_{d+1}}{a_{w_{d_n}} w_{d_n} q_{d_n+1}}\right).$$

For $d = d_0, \ldots, d_n - 1$ the error term simplifies to

$$O\left(\frac{a_{k w_d}}{a_{w_{d_n}} q_{d_n+1}}\right) = O\left(\frac{b^{(1+\varepsilon) w_{d_n} - (1-\varepsilon) k w_d}}{q_{d_n+1}}\right),$$

which after summation over $d = d_0, \ldots, d_n - 1$ gives

$$O\left(\frac{b^{(1+\varepsilon) w_{d_n}}}{q_{d_n+1}}\right) = O\left(\frac{b^{(1+\varepsilon) w_{d_n}}}{c^{w_{d_n}}}\right).$$

Because we have chosen ε such that $b^{1+\varepsilon} < c$, this error term is $o(1)$. For $d = d_n$ the error term is just

$$O\left(\frac{a_{k w_{d_n}}}{a_{w_{d_n}}}\right) = O\big(b^{(1+\varepsilon) w_{d_n} - (1-\varepsilon) k w_{d_n}}\big),$$

which is also $o(1)$ for $k \geqslant 2$. (Here we use $\varepsilon < 1/3$.) □

To illustrate the applicability of the previous theorem we present the following corollary. Essentially, we just substitute k_n by x_n, where x_n determines the final asymptotics.

Corollary 6.4. *Let $c > 1$, $\xi \in \mathbb{R}$, $\{q_d\}_{d=1}^{\infty}$ and $\{w_d\}_{d=1}^{\infty}$ be as in Proposition* 6.1. *Moreover, let*

$$a = \sum_{k \geqslant 1} a_k \boldsymbol{e}_{k\xi},$$

and assume $1 < b < c$ *such that*

$$b = \liminf_{k\to\infty} |a_k|^{-1/k} = \lim_{n\to\infty} |a_{w_{d_n}}|^{-1/w_{d_n}}.$$

Let $d_n \in \mathbb{N}$ *and* $d_n \to \infty$. *Now assume that* $x_n \in \mathbb{R}^+$ *is any sequence such that*

$$\frac{c}{b} > \lim_{n\to\infty} x_n^{1/w_{d_n}} > \frac{1}{b}. \tag{79}$$

Then $h(n) = k_n w_{d_n}$ *with*

$$k_n = \left[\frac{x_n}{w_{d_n}|a_{w_{d_n}}|}\right]$$

is a distinguished sequence, and

$$\sum_{j=0}^{h(n)-1} a(j) = x_n \frac{a_{w_{d_n}}}{|a_{w_{d_n}}|}\big(1 + o(1)\big) + o(1), \quad n \to \infty. \tag{80}$$

Proof. Because of $|a_{w_{d_n}}|^{-1/w_{d_n}} \to b$, it is easily seen that

$$k_n \sim \frac{x_n}{w_{d_n}|a_{w_{d_n}}|}$$

and that $\lim_{n\to\infty} k_n^{1/w_n}$ exists and lies between 1 and c. In particular, $k_n \to \infty$ as $n \to \infty$. Using (68) it follows that

$$\frac{k_n w_{d_n}}{q_{d_n+1}} \sim \frac{x_n}{q_{d_n+1}|a_{w_{d_n}}|} \sim \frac{x_n}{c^{w_{d_n}}|a_{w_{d_n}}|}.$$

This converges to zero because if we take the w_{d_n}-th root, the limit is less than 1. Hence the previous theorem can be applied. In particular, h is a distinguished sequence. It is now easy to conclude (80) from (77). □

Given $1 < b < c$ and ξ (and hence q_d and w_d), we can now choose x_n and d_n in different ways in order to realize several kinds of asymptotics. We have also some freedom in choosing a_k and d_n. We will choose, for simplicity, that $a_k = b^{-k}$, $k \geqslant 1$, i.e.,

$$a = \sum_{k=1}^{\infty} b^{-k} \boldsymbol{e}_{k\xi},$$

and $d_n = n$.

Example 6.5. Given $1 < b < c$ and ξ, choose $x_n = x > 0$. Obviously,

$$\lim_{n\to\infty} x^{1/w_n} = 1.$$

Hence the corollary can be applied, and we conclude that with the sequence a as above there exists a distinguished sequence h such that

$$\sum_{j=0}^{h(n)-1} a(j) = x + o(1), \quad n \to \infty.$$

Example 6.6. Choose $x_n = 1/w_n$. Since $\lim_{n\to\infty} x_n^{1/w_n} = 1$, the corollary can be applied and it follows that there exists a distinguished sequence h such that

$$\sum_{j=0}^{h(n)-1} a(j) = o(1), \quad n \to \infty.$$

This is, of course, not a counterexample, but it shows that the "desired" asymptotics holds at least for some distinguished sequences.

Example 6.7. Choose $x_n = b^{w_n\alpha/(1-\alpha)}$ with α such that

$$0 < \alpha < 1 - \log b / \log c < 1.$$

Clearly, $\lim_{n\to\infty} x_n^{1/w_n} = b^{\alpha/(1-\alpha)}$, which lies between $1/b$ and c/b because of the previous assumption. A simple computation implies that $h(n) \sim b^{w_n/(1-\alpha)}$. We can conclude that

$$\sum_{j=0}^{h(n)-1} a(j) = b^{w_n\alpha/(1-\alpha)}\bigl(1 + o(1)\bigr) = h(n)^\alpha\bigl(1 + o(1)\bigr), \quad n \to \infty.$$

We see that now this term can be even unbounded. Recall the fact that in any case, the asymptotics is $o(h(n))$ as $n \to \infty$. If we choose b and c properly in the beginning, then each $0 < \alpha < 1$ can occur as a possible exponent.

References

[1] B. Aupetit, A Primer on Spectral Theory, Springer-Verlag, New York, Berlin, Heidelberg, 1990.
[2] A. Avila, S. Jitomirskaya, The Ten Martini problem, Ann. of Math. (2) 170 (1) (2009) 303–342.
[3] A. Baker, Linear forms in the logarithms of algebraic numbers I, Mathematica 13 (1966) 204–216;
A. Baker, Linear forms in the logarithms of algebraic numbers II, Mathematica 14 (1967) 102–107;
A. Baker, Linear forms in the logarithms of algebraic numbers III, Mathematica 14 (1967) 220–228.
[4] A. Baker, Transcendental Number Theory, Cambridge University Press, London, New York, 1975.
[5] V.I. Bernik, M.M. Dodsen, Metric Diophantine Approximation on Manifolds, University Press, Cambridge, 1999.
[6] F.P. Boca, Rotation C^*-Algebras and Almost Mathieu Operators, Theta Ser. Adv. Math., vol. 1, The Theta Foundation, Bucharest, 2001.
[7] A. Böttcher, B. Silbermann, Invertibility and Asymptotics of Toeplitz Matrices, Akademie-Verlag, Berlin, 1983.
[8] A. Böttcher, B. Silbermann, Analysis of Toeplitz Operators, Akademie-Verlag, Berlin, 1989; Springer-Verlag, Berlin, Heidelberg, New York, 1990.
[9] A. Böttcher, B. Silbermann, Introduction to Large Truncated Toeplitz Matrices, Springer-Verlag, Berlin, Heidelberg, 1999.
[10] J.W.S. Cassels, An Introduction to Diophantine Approximation, University Press, Cambridge, 1957.
[11] C. Corduneanu, Almost Periodic Functions, Interscience Publishers, John Wiley & Sons, New York, 1961.

[12] T. Ehrhardt, A new algebraic approach to the Szegö–Widom limit theorem, Acta Math. Hungar. 99 (3) (2003) 233–261.
[13] N.I. Feldman, An improvement of the estimate of a linear form in the logarithms of algebraic numbers, Mat. Sb. (N.S.) 77 (119) (1968) 423–436 (in Russian).
[14] I. Gohberg, S. Goldberg, N. Krupnik, Traces and Determinants of Linear Operators, Oper. Theory Adv. Appl., vol. 116, Birkhäuser, Basel, 2000.
[15] M.G. Krein, On some new Banach algebras and theorems of Wiener–Levy type for Fourier series and integrals, Mat. Issled. 1 (1) (1966) 82–109 (in Russian); Engl. transl.: Amer. Math. Soc. Transl. Ser. 2 93 (1970) 177–199.
[16] J. Puig, Cantor spectrum for the almost Mathieu operator, Comm. Math. Phys. 244 (2004) 297–309.
[17] S. Roch, Finite sections of band-dominated operators, Mem. Amer. Math. Soc. 191 (895) (2008).
[18] S. Roch, B. Silbermann, Szegö limit theorems for operators with almost periodic diagonals, Oper. Matrices 1 (1) (2007) 1–29.
[19] K.F. Roth, Rational approximations to algebraic numbers, Mathematica 2 (1955) 1–20.
[20] W. Rudin, Functional Analysis, McGraw–Hill, Inc., New York, 1991.
[21] W.M. Schmidt, Simultaneous approximation to algebraic numbers by rationals, Acta Math. 125 (1970) 189–201.
[22] B. Simon, Orthogonal Polynomials on the Unit Circle. Part 1: Classical Theory, Amer. Math. Soc. Colloq. Publ., vol. 54, Amer. Math. Soc., Providence, RI, 2005.
[23] G. Szegö, Ein Grenzwertsatz über die Toeplitzschen Determinanten einer reellen positiven Funktion, Math. Ann. 76 (1915) 490–503.
[24] G. Szegö, On certain Hermitian forms associated with the Fourier series of a positive function, in: Festschrift Marcel Riesz, Lund, 1952, pp. 222–238.
[25] M. Waldschmidt, Diophantine Approximation on Linear Algebraic Groups, Grundlehren Math. Wiss., vol. 326, Springer-Verlag, Berlin, 2000.
[26] H. Widom, Asymptotic behavior of block Toeplitz matrices and determinants II, Adv. Math. 21 (1976) 1–29.